The Intensity of the Insight Experience in Problem Solving: Structural and Dynamic Properties

by

Natalia Derbentseva

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Management Sciences

Waterloo, Ontario, Canada, 2007

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Natalia Derbentseva

Abstract

Field theory of Lewin was used to analyze the experience of insight problem solving. It was proposed that insight is characterized by the intensity of the experience at the moment of solution. It was argued that the intensity of the insight experience depends on the experienced degree of difficulty of the problem for an individual. The experienced degree of difficulty was conceptualized as a two-fold notion: It was defined by the interdependence of the degree of restructuring involved in the problem and the dynamics of the solution process, which causes the change in the state of tension experienced by the problem solver.

Two hypotheses were formulated outlining the relationship between the intensity of the insight experience and both the degree of restructuring required to solve the problem and the amount of tension released in the system with the solution. The developed theoretical framework was investigated in the domain of matchstick arithmetic problems. A measure of the degree of restructuring for this domain was developed, and a preliminary test of the measure was carried out. Four experiments were conducted to investigate the effects of the degree of restructuring and the amount of tension on the intensity of the insight experience.

The results showed that the solution of a problem that required higher degree of restructuring resulted in a more intense experience of insight. Moreover, when the same problem was solved with higher level of tension, it led to a more intense experience of insight. Thus, it was empirically shown that the intensity of the insight experience was affected by both structural and dynamic properties of the solution process. The theoretical framework, the design of the experiments, and the results are discussed.

Acknowledgements

I would like to express my deepest gratitude to Professor Frank Safayeni, whose insightful guidance, continuous encouragement, and fascinating sense of humor made the undertaking of this work not only possible but also very enjoyable. Throughout the years of my graduate training, Professor Safayeni has shared with me his unique approach to working on research problems, an approach of critical thinking, careful examination, and the determination to get to the essence of the problem. I am endlessly indebted to Professor Safayeni for his creative ideas and his invaluable teaching.

This work would not have been possible without continuous support of my loving family, my son Sergey and my husband Ruslan, who have been tremendously caring and understanding throughout all these years. My family has always been there for me on this uneasy path.

I would like to thank my Examination Committee Members, Dr. James MacGregor, Dr. Derek Koehler, Dr. Rob Duimering, and Dr. Scott Jeffrey for their creative suggestions and all the time and effort invested into understanding this work.

I am grateful for vital contribution of four hundred UW undergraduates who participated in my studies and made the data collection possible.

I would like to thank Justin Safayeni for his work on editing and improving readability of this thesis.

Finally, I would like to express my great appreciation to Social Sciences and Humanities Research Council of Canada for financial support during the years of my PhD training.

Table of contents

Abstract ... iii

Acknowledgements ... iv

Table of contents ... v

List of Tables ... ix

List of Figures ... x

List of Figures ... x

Introduction ... 1

1. Research on insight ... 4
   - 1.1. Gestalt research on insight ... 4
   - 1.2. Information processing theories of insight ... 9
     - 1.2.1. Suddenness of insight ... 9
     - 1.2.2. Taxonomy of insight problems ... 10
     - 1.2.3. Explanations of insight ... 12
       - 1.2.3.1. Incubation stage hypotheses ... 12
       - 1.2.3.2. Search for a new representation ... 14
       - 1.2.3.3. Representational change theory ... 15
       - 1.2.3.4. Progress monitoring theory ... 17
     - 1.2.4. Sources of difficulty in insight problems ... 19
     - 1.2.5. Measuring restructuring ... 21
     - 1.2.6. Dynamic properties of insight problem solving ... 23
   - 1.3. Critical summary ... 24
2. Theoretical framework ... 26
   - 2.1. Degree of restructuring ... 27
     - 2.1.1. The amount of change in insight problems ... 28
     - 2.1.2. Quantification of restructuring proposed by Knoblich and colleagues (1999) 31
       - 2.1.2.1. Levels of problem representation ... 31
       - 2.1.2.2. Constraints in matchstick equations ... 31
       - 2.1.2.3. Mapping between constraints and levels of representation ... 32
       - 2.1.2.4. Evaluation of the approach ... 33

2.1.3. Matchstick equations and the measurement of change 34
2.1.3.1. Relationships in matchstick equations 35
2.1.3.2. Measuring change in a matchstick equation 38
2.2. Dynamic component of the insight problem solving process 42
2.2.1. Field theory of Kurt Lewin 42
2.2.2. Immediate psychological situation 44
2.2.3. Dynamics of the process of solving an insight problem 46
2.2.4. Dynamics of tension in insight problem solving 47
2.2.4.1. Structure of the life space at the beginning of the problem solving process 47
2.2.4.2. Dynamics of the process of looking for a solution 53
2.2.5. Characteristics of tension in the insight problem solving 60
2.3. Theoretical propositions...62
3. Method 65
3.1. Preliminary investigation 65
3.1.1. Method 65
3.1.2. Results...66
3.1.2.1. Experience before the solution...66
3.1.2.2. Experience at the moment of solution 68
3.1.3. Discussion 69
3.2. Experiment 1 71
3.2.1. Method 71
3.2.2. Results...75
3.2.2.1. Testing Hypothesis 1...75
3.2.2.2. Verifying the measure of restructuring 76
3.2.3. Discussion 81
3.2.3.1. Possible adjustment to the change score 81
3.2.3.2. Evaluation of the results with respect to taxonomy of matchstick problems proposed by Knoblich et al. (1999)..83
3.2.3.3. Evaluation 86
3.3. Experiment 2 88

3.3.1. Method ... 88
3.3.2. Results ... 92
3.3.3. Discussion ... 94
3.4. Experiment 3 ... 96
3.4.1. Method ... 97
3.4.2. Results ... 103
3.4.2.1. Dichotomous representation ... 104
3.4.2.2. Analysis of the relative measures ... 108
3.4.2.3. Discussion ... 111
3.4.3. Tension curve analysis: Results and discussion ... 114
3.4.3.1. Overall changes in tension before and after the solution ... 115
3.4.3.2. Tension relationship between the benchmark and the experimental problems ... 117
3.4.3.3. Pattern of tension increase ... 118
3.4.3.4. Dynamics of tension after the solution ... 122
3.4.4. Discussion ... 123
3.5. Experiment 4 ... 126
3.5.1. Method ... 128
3.5.2. Results ... 128
3.5.3. Discussion ... 132
4. General Discussion ... 134
4.1. Intensity of the insight experience ... 134
4.2. Restructuring in insight problem solving ... 135
4.3. Tension in insight problem solving ... 136
4.3.1. Effect of tension on performance ... 137
4.3.2. Tension release versus tension generation ... 141
4.4. Possible alternative explanation: concept of subjective probability ... 143
4.4.1. Subjective probability in the context of problem solving ... 143
4.4.2. Dynamics of subjective probability during problem solving ... 144
4.4.3. Psychological representation of the experience ... 145

4.5. Application of Field Theory to the study of the intensity of the insight experience 146
4.6. Limitations 147
4.7. Future research 149
Conclusions 151
References 153
Appendix A: A short summary of Lewin's theory 158
Appendix B: List of Roman numerals from 1 to 12 and mathematical operations 168
Appendix C: Descriptions of the solution moment by participants in Preliminary Investigation 169
Appendix D: Example of a form sent to the participants in Experiment 1 (reduced size) 171
Appendix E: Transcript of verbal instructions given to participants in Experiment 2 188
Appendix F: Blank tension graph used in Experiments 2, 3, and 4 (reduced size) 192
Appendix G: Example of a completed tension graph (reduced size) 193
Appendix H: Transcript of video instructions used in Experiments 3 and 4 194

List of Tables

Table 1: Heuristics and rationale for counting the change in matchstick equations.…….…..39
Table 2: Descriptor words of the experience before the solution was found..........................67
Table 3: Solution moment descriptor words..68
Table 4: Ten matchstick equations used in Experiment 1 71
Table 5: The 24 pairs of matchstick problems used in Experiment 1....................................73
Table 6: Judgment of insight for the two problems in the first comparison pair (*N*=139) 76
Table 7: Descriptive statistics for the judgment of difficulty for the problem with lower and higher change score in each pair analyzed..79
Table 8: Results of single Wilcoxon Signed Ranks tests computed for 21 pairs along with the significance assessment for each test result based of the FDR procedure at 0.05 and 0.06 levels 80
Table 9: Comparison of the experimental predictions made based on the change score versus the taxonomy of matchstick problems (Knoblich et al., 1999)...84
Table 10: Descriptive statistics for solution time (*seconds*), tension release (*cm*), intensity of insight (*cm*), *N* = 28 94
Table 11: Relative Solution Times (in seconds) in Experiment 3 109
Table 12: Means and standard deviations of the relative measures of tension, insight, and perceived difficulty 110
Table 13: Spearman correlation coefficients between the relative measures (N = 117) 111
Table 14: Changes in the tension curves (centimetres) 115
Table 15: Tension curve relationships between the benchmark (B) and the experimental (E) problems...117
Table 16: Increases in tension curves (centimetres) 118
Table 17: Solution times (second), levels of tension, intensity of insight, and perceived difficulty...129
Table 18: Results of Wilcoxon signed ranks tests for the difference between the first and the second problem 130
Table 19: Tension curve increases across three intervals before the solution 131

List of Figures

Figure 1: Classical example of an ambiguous figure: Young lady / old lady..........................28
Figure 2: Structure of a problem solving region with a Goal, Failure, and Initial State sub-regions, the person (P) is located in the initial state 50
Figure 3: Problem solving region at the beginning of the problem solving process 51
Figure 4: Insight Intensity Scale 92
Figure 5: Problem difficulty scale..101
Figure 6: Binary report of the solution times in the three groups..105
Figure 7: Binary report of the amount of tension released after the solution in the three groups 106
Figure 8: Binary report of the intensity of the insight experience in the three groups 107
Figure 9: Binary report of the difficulty judgment in the three groups 107
Figure 10: Frequencies of reported tension reduction times (in seconds) after the solution. 123

Introduction

> "What happens when a problem is solved, when one suddenly "sees the point"?"
>
> (Wertheimer, 1925/1969, p.1).

Search for an answer to this question was the motivation for the present work.

For the past eight decades researchers have continued to be fascinated with the phenomenon of insight in human problem solving. Many different questions have been raised and answers to them have been sought and proposed. These questions range from whether such phenomenon even exists (e.g. Weisberg & Alba, 1981) to what mechanisms may account for solving insight problems (e.g. Kaplan & Simon, 1990; MacGregor, Ormerod, & Chronicle, 2001). In the literature, some problems that are classified as insight are tremendously difficult as in the case of the famous nine-dot problem (Scheerer, 1963/1967; MacGregor et al., 2001; Kershaw & Ohlsson, 2004), which is practically unsolvable without any hints or additional information. On the other hand, there are much "easier" problems that are also considered to be in the category of "insight." For example, many of the matchstick arithmetic problems used by Knoblich et al. (1999) can be solved by most people within a five-minute time interval.

Different insight problems vary in their degree of difficulty. An obvious question is what makes an insight problem difficult? Gestalt psychologists proposed that the structural properties of the problem as well as the past experience of the person contribute to problem difficulty. Kershaw and Ohlsson (2004) suggested that different problems might have multiple unique causes of difficulty such as perceptual factors, prior knowledge and experience, and processing demands. Knoblich et al. (1999) suggested that the difficulty of problem solving depends on the required amount of change in the problem representation.

In this research two aspects of the problem difficulty were investigated:

- The structural properties were conceptualized and quantified as the degree of restructuring in the problem representation;
- The dynamic component of the person's experience, which has not been investigated by researchers, was also considered to contribute to the experience of problem difficulty. Lewin's field theory was used to analyze the dynamics of solving an insight problem.

The question that has not been asked yet is how the different degrees of difficulty affect the experience of insight. Do different insight problems lead to the same experience of insight? Posing this question differently: Does a simple insight problem and a very difficult insight problem result in the same experience of insight? In other words, is there a dimension of insight, namely an intensity that varies depending on the problem? In this thesis, it was hypothesized that different insight problems may result in different degrees of insight.

Solving an insight problem is experienced in the context of encountering impasses, getting stuck, being uncertain as to what to do next, sometimes even abandoning the process altogether, and, after a struggle, finally, finding the solution and experiencing the "Aha!" Researchers in the field of insight problem solving agree that impasse is a hallmark of the insight problem solving process, although not necessary for the insight to occur. The dynamic aspects of the process of problem solving have been largely neglected by cognitive psychologists. In the study of insight, the emphasis has been on the structural properties of the problem and its search space. The proposed models of insight process (e.g. MacGregor et al. 2001; Kaplan & Simon, 1990) treat the impasse merely as an indicator that the search space needs to be expanded.

Gestalt psychologists conceived of insight as an "Aha!" experience associated with the release of tension (Ormerod, MacGregor, & Chronicle, 2002). What exactly is this tension? Where does it come from and how does it change? What role does it play in the experience of the problem difficulty and the intensity of the insight? To answer these questions, a closer look into the dynamics of the problem solving experience was undertaken. The concept of psychological tension (Lewin, 1935, 1936) was applied to understand the underlying dynamics of the insight problem solving.

This work is organized in the following manner. First, the literature pertaining to the study of insight problem solving is reviewed. Second, the theoretical framework and hypotheses are developed. Third, a preliminary study and four experiments are reported. Finally, the general implications of this work are discussed.

1. Research on insight

In the first half of the twentieth century Gestalt psychologists triggered a surge of interest in the study of the phenomenon of insight. The phenomenon of insight has received a substantial attention in the literature. The following section provides a brief summary of the research on insight in problem solving, starting from the description of Gestalt studies of insight, followed by the description of current approaches.

1.1. Gestalt research on insight

The research on insight was popularized by Gestalt psychologists in the first half of the twentieth century. Such researchers as M. Wertheimer, W. Kohler, K. Dunker, N. R. F. Maier, A. Luchins from the Gestalt school contributed to the investigation of the insight phenomenon.

Gestalt psychologists argued that any situation exists for an individual as an organized whole, termed a "Gestalt." The elements and their interrelationships in such organized whole constitute a structure. The problem-solving situation is a situation where at the start, an individual holds a whole-view of the situation that is unsuitable for the problem at hand, and, thus, initially prevents the solution (Wertheimer, 1959). Problem solving activity consists of grasping the inner structure of the situation "according to the requirements of the problem," forming a new Gestalt of the situation. The process of restructuring in problem solving constitutes a switch from one whole-view structure of the situation, not suitable for the task, to a different one. Karl Dunker described restructuring as a process by which "parts of the situation which were formerly separated as parts of different wholes, or had no specific relation although parts of the same whole, may be united in one new whole" (Dunker, 1945, p. 29). Summarizing Gestalt view of restructuring, Ohlsson (1984) wrote "restructuring is a change which affects the structural relations in the situation" (p.68). Gestalt psychologists conceptualized the phenomenon of insight as an act of restructuring of the problem solving situation which happens suddenly.

The concept of restructuring is central to the Gestalt view of insight and it was explored in Gestalt psychologists' investigations. One of the earliest and the most cited work on insight problem solving is the series of experiments described by Kohler (1925) on the intelligent behaviour of chimpanzees (Mayer, 1995). In these studies, hungry animals faced the problem of getting their food that could not be obtained by the usual, most direct approach. For example, bananas were placed outside a chimpanzee's cage and could be reached only by connecting together two short sticks ; or food was hung from the ceiling of a chimpanzee's cage too high for a grab, and required stacking of a few boxes (lying around the cage) on each other and climbing to the top one. Getting the food required the animals to abandon the initial, most direct approach and find a "detour." Kohler described chimpanzees' solutions to these problems as representing "complete solutions," which were obtained suddenly and were a result of the "reorganization of the field."

Wertheimer (1959) provided another classic account of restructuring when he analysed the restructuring involved in the discovery of the sum of the series in a young Gauss story. A sum of a series of numbers (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10) has to be found without sequential addition of all the numbers. Wertheimer used different variations of this problem in his investigation with children and adults, finding that some were able to solve the problem while others failed or even refused to try. The required restructuring in this problem is to view the series not as an increasing sequence of numbers but as a collection of pairs of numbers that add up to the same quantity (e.g. 1 and 10, 2 and 9, 3 and 8).

Gestalt psychologists not only emphasised the importance of restructuring in problem solving but also investigated the nature of the difficulty that arises when restructuring introduces such notions as fixation, Einstellung, and reproductive thinking. Scheerer (1963/1967) described fixation as a process whereby "sometimes a person clings misguidedly to a false premise or assumption concerning the task before him" (p.29). For example, people often failed to construct four equilateral triangles with six matches lying on the table, because they tried to find the solution in a single plane – two-dimensional space (Scheerer, 1963/1967). Gestalt psychologists saw fixation as a force that directed problem solving and provided resistance to a restructuring.

A form of fixation studied by Dunker was termed functional fixedness. According to this concept,

> an object with a strong customary function will not easily be seen as serving a different function. More specifically, problem presentations or prior experiences that emphasise the usual function were predicted to inhibit the use of an object in a more novel fashion (Dominowski & Bourne , 1994, p. 26-27).

In the classic example of Dunker's candle holder problem, a participant was given a set of objects which included a candle, a box of tacks, and a box of matchsticks. The participant was asked to mount the candle onto the wall so that when the candle was lit the wax would not drip onto the floor (Dunker, 1945). The solution to this problem is to attach a box to the wall using tacks and attach the candle to the box by dripping a few drops of wax into the box. In one condition of the study, participants received boxes filled with objects (e.g. tacks and matches); in another condition , participants received boxes and objects separately, that is the boxes were presented empty. The solution rate was much lower in the condition where boxes were presented filled with objects. According to Dunker, the box filled with objects was imbedded in a particular context, a "functional whole" that made it difficult to separate the elements. In other words, participants were fixated on the function of the box as a container, making it difficult for them to conceptualise it as a candle holder.

The same effect was observed by Scheerer (1963/1967) in the experiment in which participants had to connect two sticks to each other, a task which could have been done by using a string that was in the room. When the string was simply hanging on the nail, the participants used it with no difficulty. However, when the string was attached to a picture that hung on the wall, the participants failed to avail of the string. The effect of fixation has been investigated beyond insight and problem-solving domain, for example in a broader context of creativity (e.g. Smith, Ward & Schumacher, 1993) and engineering design activities (e.g. Chrysikou & Weisberg, 2006).

Gestalt psychologists introduced another phenomenon that hinders restructuring – Einstellung or a mental set. Einstellung refers to applying a previously learned rule or procedure to a task when there exists a simpler way of doing the task. Luchins (e.g. Luchins & Luchins, 1950) demonstrated the set effect on a series of water jar problems. Participants were presented with a set of 10 water jar problems in which they were asked to measure a certain quantity of liquid using three measuring containers of different volume. The initial six problems had to be solved by one method that was similar for all of the initial six problems. This initial set of six problems was called the Einstellung set. Although the remaining problems could also be solved via the initial method, a much simpler way existed for them. One of the problems in the set could be solved only by a simpler method but not the initially learned one. People who solved the Einstellung set tended to continue using the same method on the remaining problems that had simpler solutions. Moreover, many of these participants failed to solve the problem that required a simpler solution and that could not have been solved by the learned method.

McGraw and McCullers (1979) used Luchins and Luchins (1950) water jar problems to create the mental set effect in participants of their study. The authors demonstrated that the participants who were paid for correctly solving the problems had more difficulty overcoming the fixation effect (measured by solution time) than those who did not receive the payment (McGraw & McCullers, 1979). Einstellung is a negative effect of past experience which Luchins described as "instead of the individual mastering the habit, the habit masters the individual" (Luchins & Luchins, 1950, p. 279).

Ohlsson (1992) indicated that the Einstellung effect was not tested for other task domains than the water jar problems. Ohlsson (1992) also argued that Einstellung is a different phenomenon, not related to insight, since there are no impasses involved. Indeed, the impasse will not be reached if the learned method leads to the solution of a new task. However, Luchins showed that when the learned method was inappropriate for the problem at hand, it led to great difficulty in solving the problem, with participants even failing to find the solution altogether. What could be a better manifestation of reaching a state of impasse than a definite failure to solve the problem? It is worth noting that the same problem was solved with relative ease when the participants did not work on the Einstellung set prior to attempting the problem. The Einstellung is a "blinding effect" of habit that prevents people from productive restructuring of the situation.

In his work "Productive thinking," Wertheimer discussed the difference between "productive" thinking as requiring understanding of the interrelationships in the situation and "reproductive" thinking as the "blind" application of previously learned approach (Wertheimer, 1959). This distinction was illustrated with the example of finding the area of a parallelogram. Wertheimer showed that students who merely memorised the sequence of operations in the procedure failed to transfer their knowledge to a rotated version of the figure. Katona (1940, cited in Mayer, 1995) demonstrated similar results in the study on the solution of a matchstick puzzle. Reproductive thinking, according to Wertheimer, is structurally blind (Wertheimer, 1959).

Gestalt psychologists were criticised for proposing a theory of "good" thinking only without paying attention to "not good" thinking, in that they provided vague explanations that lacked experimental support and methodological rigor (e.g., Ohlsson, 1984; Weisberg & Alba, 1981). While these criticism are valid for the most part, one needs to keep in mind the historical environment in which Gestalt theory was formulated. Dominowski and Bourne (1994) argued that "Gestalt theory, although its principles were sometimes rather vague, helped psychology to achieve a more balanced and realistic view of complex human behaviour than would have been possible on the basis of Behaviorism alone" (Dominowski & Bourne , 1994, p. 18). As Mayer (1995) put it, Gestalt psychologists "left us with some questions and the beginnings of some answers" (p. 26). The phenomenon of insight is one of such questions left to us by the Gestalt school.

Inspired by the Gestalt psychologists, contemporary scholars have undertaken the investigation of insight problem solving, trying to find explanations of this phenomenon. The investigation of insight still remains an active topic in the literature. The following section will summarise the significant recent developments in the research on insight.

1.2. Information processing theories of insight

The Gestalt notion of "restructuring" is interpreted in modern cognitive psychology as a "process which changes the problem solver's mental representation of the problem" (Ohlsson, 1984, p.71) and remains central to the investigation of the insight phenomenon (e.g. Gilhooly & Murphy, 2005). Information processing models of insight account for the Gestalt notion of restructuring through standard information processing mechanisms (e.g. Kaplan & Simon, 1990). "Under their models, the restructuring of a problematic mental representation stems from retrieval processes that search semantic memory for relevant concepts. Difficulties associated with reaching a successful solution are then attributed to a failure in accessing the right solution plans from memory" (Seifert et al., 1995, p.71). However, both characteristics of the insight experience proposed by the Gestalt school – restructuring and suddenness - are present in the current investigations of this phenomenon.

1.2.1. Suddenness of insight

The phenomenological property of suddenness of insight problem solving was demonstrated in the studies of Janet Metcalfe (Metcalfe, 1986a, 1986b; Metcalfe and Wiebe, 1987). In the first study, Metcalfe (1986a) demonstrated that the solution to the insight problem could not be predicted in advance contrary to the memory questions. In the second study, Metcalfe (1986b) showed the suddenness characteristic of insight solutions using the feeling-of-warmth ratings. Every 10 seconds the participants were asked to estimate how close (i.e. warm) they thought they were to arriving at a problem solution while working on the insight problems. The results showed that the feeling-of-warmth ratings remained low on insight problems up until the solution was found. In the third study, Metcalfe and Wiebe (1987) demonstrated the difference between insight and non-insight problem solving processes using feeling-of-warmth ratings, also. These ratings were assessed every 15 seconds during the solutions of insight and non-insight problems. The results showed that an incremental increase in feeling-of-warmth ratings characterized non-insight problems, and that these ratings were good predictors of performance. The warmth ratings for insight problems "showed a more sudden achievement of solution."

Ohlsson (1992) argued that the solution of insight problems is experienced as sudden because of the breaking out of impasse through forming a correct problem representation. Similarly, Gick and Lockhart (1995) suggested that suddenness of insight derives from the new representation leading to an immediate solution of the problem. These authors also suggested that the affective response of surprise and suddenness when an insight problem is solved is due to the difference between the original and solution representations of the problem. Furthermore, "the cognitive components of insight are not necessarily sudden, but the accompanying affective components do have the quality of suddenness and surprise" (Gick & Lockhart, 1995, p. 215).

1.2.2. Taxonomy of insight problems

Weisberg (1995) accurately noted that the field of studying insight problem solving suffers from the lack of a clear and commonly accepted definition of what constitutes an insight problem. Some researchers proposed to differentiate between insight and non-insight problems based on the suddenness of obtaining the solution. Metcalfe & Wiebe (1987) proposed that insight problems are characterized by a sudden discovery of a solution but that non-insight problem solving requires an incremental approach. However, this method allows only a post hoc differentiation of the problems; without some independent basis for classification, suddenness of obtaining the solution becomes both the basis for classification and a support for it (Weisberg, 1992).

Schooler et al. (1995) suggested that the difference between insight and non-insight problems
is a discrimination between approach-recognition (insight problems) versus approach-execution (non-insight problems). Grant and Spivey (2003) suggested that insight problems are the problems for which solutions cannot be logically induced (in contrast to non-insight problems, like those of algebra). However, neither group of authors suggested how problems could be classified on the proposed dimensions.

Davidson (1995, 2003) discussed the classifying of insight problems on the basis of the required solution process. The author suggested that there are three types of insight processes: selective encoding, selective combination, and selective comparison. A restructuring of the problem representation is achieved via these processes. Selective encoding refers to the process of encoding information that was previously considered irrelevant to the problem. Selective combination refers to the process of recombining elements of the problem that were encoded. Selective comparison refers to the process of relating previously learned information to the problem at hand. The three processes seem to overlap and the criteria for classification were not specified. It also remains unclear whether the problems are actually solved via the suggested processes (Weisberg, 1995).

Weisberg (1995) proposed to differentiate insight problems from non-insight problems, given the Gestalt notion of restructuring. According to Weisberg, a pure insight problem must involve discontinuity in thinking that is brought on by restructuring – that is change in the problem solver's representation of the problem. The representation of the problem consists of problem elements, relationships among them, operators that are available, and the goal of the problem. A pure insight problem can only be solved via restructuring. If another way of solving the problem is available, then the problem is classified as a hybrid. The suggested taxonomy based on the restructuring involved in a problem's solution is a useful way to identify insight problems from non-insight problems. Weisberg (1995) proposed a way of identifying restructuring as a change in elements, and/or their relationships, and/or operators, and/or goal of the problem; however, he did not elaborate further. Gilhooly and Murphy (2005) used Weisberg's (1995) taxonomy of insight problems to test for the difference between insight and non-insight problems. Their results generally supported Weisberg's classification.

It is worth noting that in Weisberg's (1995) classification of insight problem, the concept of restructuring was treated rather as a binary category – a problem is either solved via restructuring or not so solved.

Chronicle, MacGregor, and Ormerod (2004) proposed another tentative criterion by which insight problems can be distinguished: a "solution-recoding" hypothesis, which refers to the ability to reproduce the solution "without extensive search for moves" at later points in time. According to Chronicle et al. (2004), a solution to a problem that can be described as a "single executable concept" can be easily reproduced – as they found was the case for the 10-coin problem. These authors pointed out that the solution-recoding hypothesis is tentative and needs further exploration and testing; however, a good retention of the solution to an insight problem had been previously demonstrated. In Kohler's studies, apes were able to repeat their solutions when they encountered the same problem again (Kohler, 1925); subjects that solved the problems themselves "showed a near perfect memory for the solution" one week later, contrary to the subjects who were shown the solution (Dominowski & Dallob, 1995).

1.2.3. Explanations of insight

Several approaches have been adopted to explain insight phenomena by providing an account of why impasse occurs in these problems and how it is overcome (e.g. Knoblich et al., 2001).

1.2.3.1. Incubation stage hypotheses

To answer a question on how insight is facilitated by a period of interruption in the problem solving process, the phenomenon of insight is expressed in the context of a four-stage thinking process proposed by Wallas (1926, cited in Mayer, 1983). The four stages are preparation, incubation, illumination, and verification. The preparation stage consists of gathering information and attempting to solve a problem. At this stage, a problem's representation is formed, and initial attempts at solving it fail. During the incubation stage, the problem solver puts the problem aside and attends to other things. In the illumination stage, the key to the solution appears in the person's mind, and it is verified during the evaluation stage.

Seifert et al. (1995) summarised several proposed hypotheses of the function of the incubation stage or processes that take place that are beneficial for the problem solving. The conscious-work hypothesis suggests that during an incubation phase, covert conscious work on the problem is performed. The fatigue-dissipation hypothesis suggests that the incubation phase gives solver a chance to rest from the previous preparation work. The selective-forgetting hypothesis suggests that the incubation phase allows the forgetting of inappropriate solution strategies. The subconscious random-combination hypothesis suggests that during the incubation stage, segments of information are randomly recombined subconsciously, leading perhaps to a correct combination. Seifert et al. (1995) proposed an opportunistic-assimilation hypothesis of the incubation phase, which states that failure indices are stored in the long-term memory upon reaching an impasse. During the incubation stage no further work on the problem is carried out, but when the person encounters environmental stimuli relevant to the problem, the failure indices trigger a process through which the new stimuli are assimilated into a previous memory representation.

Smith (1995) proposed the “mental ruts” hypothesis, which states that repeated exploration of an unsuccessful path adds more memory activation to this path. During the incubation stage, memory traces to the blocking responses become weaker. This trend improves the likelihood of eliciting the correct response upon resumption of the problem-solving activity. According to Smith (1995), the benefit of incubation is the strengthening of the relative accessibility of the target information in the memory. This advantage depends on the duration of the incubation stage as well as on the contextual change.

Segal (2004) proposed another incubation phase hypothesis: the returning-act hypothesis. According to this hypothesis, no activity occurs during the incubation phase, and the incubation does not depend on the external environment, while the only function of a break is to “divert the attention of the solver from the problem, thus reducing or erasing the activation of the false assumption” (p.143). When the problem is re-encountered, the problem elements can assume a new organization. As can be seen, there is great similarity between the explanation proposed by Smith (1995) and that by Segal (2004). While Smith emphasises the strength of memory traces, Segal emphasises the organization of the problem, which weakens during incubation and can more easily assume a new organisation upon resumption.

Experimental evidence of benefits of the incubation stage was not always found (Segal, 2004, Seifert et al., 1995). In the laboratory investigation of the effects and processes of the incubation phase, participants are usually presented with a problem – or memory question. Then, sometime during the solution process, the participant is interrupted, and sometime later is given a chance to continue the problem solving. The duration of the preparation stage is varied, that is duration of time before interruption. The duration of the incubation stage is also varied, that is for how long participants are interrupted. The activities that participants are involved in during the incubation phase are also varied. The solution rates are measured and compared across the different preparation/incubation conditions (e.g. Segal, 2004; Seifert et al., 1995). Even though the reported evidence supports a certain hypothesis, Seifert et al. (1995) correctly point out that various conclusions might be reached based on the same results from the incubation studies, making it hard to identify the correctness of any of the explanations.

The incubation hypotheses are concerned with what happens when a problem solver stops deliberate work on the problem and engages in other activities. The proposed explanations of the incubation stage of the insight problem solving do not discuss issues related to problem difficulty. These explanations are hypotheses of cognitive mechanisms (e.g. forgetting, combining elements, encountering relevant information, etc.) without including the dynamics of the experience of an individual. Nor are these explanations useful for the situations where there is no evident incubation stage – when a problem is solved in a persistent, continuous attempt that results in success.

1.2.3.2. Search for a new representation

Kaplan and Simon (1990) proposed a process theory of insight as a search for a new problem representation (problem space), suggesting that the same processes that are used for a search within a problem representation are used to search for a problem representation. Authors have pointed out that subjects "almost always adopt the representation suggested by the verbal problem statement." Thus, a choice among problem representations is not an issue until an impasse is encountered. An impasse is reached when no operator available within the current representation yields progress toward the goal. Another necessary condition for a new representation search is the “attention to cues that guide the generation of alternative problem space.” The search space among alternative problem representations would be too huge, and impossible to operate without certain constraints that guide the search process. Constraints limit the possibilities that the subject considers and make the search manageable. The discussed sources of constraints that guide the search for a new representation in a mutilated checkerboard problem were

- Features of the problem (e.g. salience of the features)
- Hints from the experimenter (e.g. that covering the board with dominoes does not lead to a solution)
- Relevant domain knowledge (e.g. mathematical background)
- Use of heuristics (e.g. noticing invariants, comparing different board situations)

Kaplan and Simon proposed that the most interesting results of their study concerned the noticing of invariant heuristic: that is, attending to invariant features, such as two squares that are removed are always of the same color. They suggested that noticing invariant heuristic “can facilitate insight across a wide variety of domains” (p.413). However, it is not clear to what extent the invariant heuristic is applicable to other insight problems – for example, two-string problem, radiation problem, or matchstick arithmetic problems (Knoblich et al., 2001).

Kaplan and Simon (1990) discussed the possible constraints operating in this search process (i.e. information that is helpful) but not the process of finding the new representation per se, thus, leaving the dynamics of the experience out of the picture. This theory also does not address the issue of the difference between the representations.

1.2.3.3. Representational change theory

Ohlsson (1984b, 1992) proposed an alternative theory of representational change. In his information processing theory of insight, Ohlsson (1984b, 1992) conceptualized insight as “the act of breaking out of impasse” (p. 4), which occurs as a result of change in problem representation. Ohlsson viewed impasse as a necessary condition for insight, and explained suddenness of insight as a momentary shift from the state of not knowing the solution to the state of knowing it. A problem solver initially forms a representation of a problem that does not lead to the solution, because relevant operators were not retrieved. As a result, an impasse is encountered. Ohlsson suggested that “the impasse is broken by seeing the problem in a new way, just as the Gestalt psychologists claimed” (p.12). To overcome an impasse, a problem solver has to re-perceive the problem and form a new problem representation that might lead to a solution.

Ohlsson suggested three mechanisms for change in a problem representation that might lead to a solution: elaboration, re-encoding, and constraint relaxation. The incomplete problem representation can be changed by adding information that the problem solver had not noticed before (elaboration). For example, in a mutilated checkerboard problem (Kaplan & Simon, 1990), noticing parity principle leads to elaboration of the problem representation (Ohlsson, 1992). When a problem representation is mistaken, the solver has to “re-encode” the representation. For example, “seeing” a pliers as a pendulum weight in Maier’s two-string problem leads to re-encoding of the problem representation. Constraint relaxation refers to the relaxing of self-imposed constraints on the goal representation of the problem. In a six-match problem, for example, people often impose a constraint – that the solution be in two-dimensional space (Scheerer, 1963/1967).

The proposed three mechanisms seem to be highly overlapping since all three examples provided above could be framed to fit any of the descriptions. For example, “seeing” pliers as a pendulum weight could be referred to as relaxing the constraint of seeing pliers as a “grabbing device”; or it could be explained as elaboration of the problem representation by adding information about the weight property of the pliers. It is worth noting also that in order to construct a new problem representation, it is not enough just to relax a certain constraint. When a new problem representation is formed, besides relaxing an existing constraint, a new set of constraints has to be imposed. In a six-match problem, for example, relaxation of a two-dimensional representation implies a new constraint on the solution – that it be constructed in three-dimensional space (in this case the only other option), otherwise a new representation cannot be achieved. The notion of forming a new representation refers to a switch from one set of constraints that define an initial representation to a new set of constraints that define a new representation. The relationship between the constraints that are relaxed and those that are imposed is not entirely clear; it needs to be explored further.

The constraint relaxation mechanism of Ohlsson's theory of representational change of insight problem solving was experimentally tested in the domain of single-move matchstick equation problems (Knoblich et al., 1999, 2001). Knoblich et al. (1999) distinguished among four types of matchstick equations based on the type of a constraint involved in their solution. The hierarchy of constraints was suggested based on the degree of effect on the problem representation. Knoblich et al. (1999) hypothesized that the difficulty of a problem depends on the scope of problem representation change required by relaxing a certain constraint. Authors reported experimental results that supported their hypotheses. Thus, Knoblich et al. (1999) introduced the idea of relationship between the amount of change in the problem representation and the problem difficulty. The notion of varying amount of change in the problem's representation was developed further in the present work as the degree of restructuring involved in insight problems. The quantification used by Knoblich and colleagues is detailed in section 2.1.2.

1.2.3.4. Progress monitoring theory

MacGregor, Ormerod, and Chronicle (2001) proposed an information processing model to explain the difficulty of solving the nine-dot problem. The model suggests that problem solvers apply "locally rational" operators at every move on the basis of some criterion of progress. The "locally rational" operator reduces the distance to the goal or current sub-goal. Among the alternative operators, that operator is selected which reduces the distance to the goal the most. The criterion of progress on the nine-dot problem could be a minimum number of dots that a line must intersect, which would depend on the number of dots left to cancel and the number of available lines. If no operators can be found that satisfies the criterion of progress, then "the problem space expands and a search for alternative operators takes place" (p. 177).

The search for alternative operators proceeds through relaxing constraints (Ohlsson, 1992). One possible constraint operating in the nine-dot problem could be the defining of lines through adjacent dots. The authors suggested that an "impulse" to look for emergent moves is proportional to the discrepancy between the current criterion and the best available operator. Individual problem solvers could have different strategies or memory capacities, which would affect their look-ahead. Look-ahead determines how many lines for a nine-dot problem an individual considers in a single move (up to four for this problem). The model's predictions were tested in the laboratory: MacGregor et al. (2001) reported results of five experiments on the solution of the nine-dot problem and its variants that supported the model's predictions. Chronicle, Ormerod, and MacGregor (2001) varied available perceptual cues available in the nine-dot problem variants; they also reported support for the progress monitoring model predictions. Some of the model's predictions were generalized to the eight-coin problem by Ormerod et al. (2002).

MacGregor et al. (2001) pointed out that application of their model is limited to multi-step problems in which a monitoring of progress is possible throughout the solution process. Thus, this model is not suitable for a domain of single-step problems such as matchstick arithmetic problems.

Jones (2003) compared predictions of the constraint relaxation theory to that of the progress monitoring theory on a novel problem, the car park game. Using eye movement data, solution times, and rates, the author tested predictions derived from the two theories. For example, from the perspective of the representational change theory, the only constraint that had to be relaxed in a car park game was that the taxi car had to be moved before the exit path was cleared. Assuming that this is the constraint that needs to be relaxed, Jones (2003) predicted that most subjects would encounter impasse before attempting to move the taxi car, and this prediction was supported by the collected eye movement data. Another prediction derived from the representational change theory was that those people who did not solve the problem should not move the taxi car. However, five out of eight non-solvers did move the taxi car, two times each on average, and still failed to solve the problem. From the progress monitoring theory, Jones (2003) made prediction with respect to the proportion of subjects with mental look-ahead of one, two, and three moves. Subjects were assigned to certain look-ahead values based on the location of the majority of their impasses in the solution process. The eye movement data supported the predicted proportion. In addition, the prediction was made with respect to performance of subjects assigned to different look-ahead values, an approach supported by the solution rate, solution time, and number of moves data. Although Jones (2003) reported that the results favoured the representational change theory when opposing predictions were generated from both theories, the accuracy of opposing predictions proposed and tested by Jones are questionable.

1.2.4. Sources of difficulty in insight problems

Kershaw & Ohlsson (2004) proposed a multiple-cause difficulty of the nine-dot problem . The causes are perceptual factors (e.g. good Gestalt, figure-ground), prior knowledge and experience (e.g. experience with connect-the-dots activities, training), and processing demands (e.g. amount of look-ahead). The authors explained previous failures to improve performance to a desired 100% level on this problem by concentrating on a single source of difficulty. The difficulty of the nine-dot problem proposed by the Gestalt psychologists is in the requirement of drawing the lines outside the strong configuration of a square formed by the dots. However, explicitly stating to the subjects that lines have to be drawn outside the dots did not result in the desired 100% solution rate that would signify that this is a sole source of difficulty (Weisberg and Alba, 1981). Kershaw & Ohlsson (2004) suggested that, most likely, all insight problems are subjects not to one but multiple causes of difficulty which could be unique for each problem.

Many scholars agree that the general source of difficulty of insight problems is related to the issue of problem representation (e.g. Kaplan & Simon, 1990; Gick & Lockhart, 1995; Ohlsson, 1992). The difficulty of changing to a correct problem representation could be due to inappropriate constraints that are imposed on the basis of prior knowledge and experience (e.g. Ohlsson, 1984b, 1992, Knoblich et al., 1999).

MacGregor, Ormerod, and Chronicle proposed in a series of studies that the difficulty of the insight problems could be due to the tendency of problem solvers to use "locally rational" operators that maximise apparent progress towards the goal (e.g. MacGregor et al., 2001; Chronicle et al., 2001). The selection of the "locally rational" operator with insufficient look-ahead might fail to succeed in insight problems that are often characterised by a "detour" property in which an immediate, most direct path to the goal does not lead to a solution.

Ormerod et al. (2002) demonstrated that the size of the available search space contributed to the difficulty of the 8-coin insight problem since the "criterion of failure" can be reached sooner when the search space is small. Ash and Wiley (2006) manipulated the size of the search space on a number of insight problems. However, these authors reported that the overall difficulty was similar on the problems with large and small search spaces.

Another potential source of difficulty in insight problem solving could be attributed to motivational factors: Several studies have reported a change in performance on insight problems resulting from the use of incentives (Glucksberg, 1962; McGraw & McCullers, 1979; Wieth & Burns, 2006).

Glucksberg (1962) reported that when participants were paid depending on how quickly they solved the Dunker's candle problem, the participants tended to take longer to solve the problem compared to the group that was not paid. Similarly, McGraw and McCullers (1979) reported that solution times on the water jar problem that required breaking a mental set increased in the group who was paid for solving the problem correctly. Camerer and Hogarth (1999) suggested that incentives in the Glucksberg (1962) and McGraw and McCullers (1979) studies led to participants' exerting more effort while working on those problems; however, "more effort blinds them to the surprising answer" (p. 22). Camerer and

Hogarth (1999) suggested that increased effort in insight problem solving could have led to persistence with one approach, making the problems more difficult.

Contrary to these findings, Wieth and Burns (2006) reported an increase in the solution rate of the insight problems in the group of participants who were offered an incentive (to leave the experiment earlier). Furthermore, these authors suggested that incentives in their studies led to a "more thorough processing," based on better remembering of the problems in the incentive condition that was assessed after solution of the fourth problem. While the solution rates in the Wieth and Burns (2006) studies increased with the use of an incentive, the solution times also increased when the incentive was used – by a small although statistically significant increment. While these authors emphasized solution rates as the appropriate measure of performance, Gilhooly and Murphy (2005) argued that the solution time is a "more discriminating" measure of performance than the solution rates. Thus, depending on which measure of performance is chosen, the findings from Wieth and Burns (2006) studies can be interpreted differently.

These studies showed that the motivational aspect of insight problem solving could also affect the difficulty of insight problems. Although the research in this area is very limited and reports somewhat contradictory results, it rightly points to other potential sources of difficulty in insight problem solving that lie outside the problem and its structure.

1.2.5. Measuring restructuring

While many researchers see the notion of restructuring as the key element of the phenomenon of insight (e.g. Ohlsson, 1992; Knoblich et al., 1999; Weisberg, 1995), Seifert et al. (1995) argued that the restructuring is not a prerequisite for an insight. Seifert et al. (1995) suggested that "restructuring constitutes one potential basis of insight, but that insight can (and does) seem as well from the addition of missing pieces to a formerly incomplete yet appropriate mental representation" (p.67). One can argue that "addition of missing pieces" does in fact restructure the problem representation since the missing pieces need to be incorporated into the previous representation, altering it and changing the interrelationships among the elements; however, Seifert et al. (1995) do not classify it as a restructuring.

Even when restructuring is considered necessary for the insight to occur, it is often treated as a rather binary category, that is, the restructuring either happens or does not. On the other hand, it might be more practical to treat restructuring as a continuum and differentiate among different 'degrees' of restructuring – different magnitudes of change to the problem representation.

The literature reports some attempts to develop a measure of restructuring (e.g. Knoblich et al., 1999; Ash and Wiley, 2004). Knoblich et al. (1999) suggested that different moves in the domain of matchstick problems cause different amounts of change to a problem representation. These authors proposed a hierarchy of problems based on the required amount of change to the problem representation. The authors also predicted the relative difficulty of categorizing problems. The quantification of restructuring proposed by Knoblich and colleagues is discussed in detail in section 2.1.2.

Ash and Wiley (2004) used a hindsight bias as a measure of the amount of restructuring involved in solving insight and non-insight problems. "Hindsight bias is the observation that people with outcome knowledge of a situation falsely believe that they would have predicted the correct outcome" (Ash & Wiley, 2004, p.1). These authors hypothesized that processes that motivate restructuring (i.e. insight problems) will lead to hindsight bias, while the problems that are not solved via restructuring (i.e. algebra problems) will not result in hindsight bias. Problems were broken down into components, and subjects were asked to rate twice the importance of problems' components. The first rating was obtained right after reading the problem for the first time before attempting to solve it. Some subjects were shown the solutions after their initial solution attempts. The second rating was obtained one week later.

The authors predicted that participants who received the answers to the problems would exhibit a significant change in the importance ratings of problems' components. This prediction was supported for insight problems. Hindsight bias is a novel way of looking at restructuring; however, it has certain limitations in its application. Using hindsight bias does not allow the creation of a continuum of restructuring and measuring different amounts of restructuring. This measure is also not accurate because the rating was done only on the elements of the problem, neglecting their interrelationships, available operators, and the goal definition (Weisberg, 1995). Rating of all of the problem components might be simply impossible.

1.2.6. Dynamic properties of insight problem solving

As in other human activities, solving an insight problem is a process characterised by certain dynamics. However, very little work has been done to investigate the dynamics of insight problem solving (Ormerod et al., 2002). MacGregor et al.'s (2001) progress monitoring model incorporated a dynamic component of search for a solution that is directed by local optimization and progress monitoring in insight problem solving. Their model, however, emphasises the information-processing aspect of insight problem solving, leaving out the aspect of the solution process that involves the dynamics of tension and getting frustrated with the problem. Until fairly recent years, an information-processing perspective generally omitted the aspects of motivation and affective response from the analysis of cognitive process (Weith & Burns, 2006).

The studies by Glucksberg (1962), McGraw and McCullers (1979), and Wieth and Burns (2006) reported different effects of motivational factors on insight problem solving performance. Glucksberg (1962) argued that the increased motivation prolongs extinction of the dominant habit (i.e. seeing the box as a container) which is not the correct response for the problem, and that it prevents the correct habit (i.e. seeing the box as a platform) from "gaining ascendancy." Camerer and Hogarth (1999) similarly suggested that increased effort in insight problem solving might lead to persistence with one approach, making the problems more difficult. Although this might be a plausible explanation of the results, there was no evidence provided in the studies that actually demonstrated the persistence with one approach during the solution process over trying a variety of approaches. Wieth and Burns (2006) reported the same effect of incentive on performance in both insight and incremental problems, which led the authors to dispute the validity of explanation by persistence with one approach. The mechanisms by which motivational factors affect insight problem solving process remain unclear.

The dynamics of the insight problem solving depend, on one hand, on the problem itself and its structure, and, on the other hand, on the state of the individual and the problem solving environment. These aspects will be discussed in the theoretical section which follows.

1.3. Critical summary

The foregoing review of the literature shows that the basic Gestalt definition of the phenomenon of insight as sudden restructuring is still very much active. Some attempts have been made to develop a taxonomy of insight problems (e,g. Weisberg, 1995); however, there is still no agreed-upon definition of what constitutes an insight problem and what processes underpin the solving of insight problems. This ambiguity creates a major difficulty for the field (Bowden et al., 2005).

Several approaches (e.g. Kaplan & Simon, 1990; MacGregor et al., 2001; Knoblich et al., 1999) have been proposed to explain why impasses are encountered, how they are overcome, and the search mechanisms of finding the insight solution. However, these approaches have been concerned primarily with the structural properties of the search space, problem representation, search process, and strategy, all largely ignoring the dynamics of problem solver's experience.

Bowden et al. (2005) pointed out that different theories of insight address different components of this phenomenon, making the comparison of theories' predictions difficult.

Scholars of insight largely agree that a central characteristic of insight is the notion of restructuring. We still do not have, however, a reasonable measure of restructuring applicable to at least a set of homogeneous insight problems. The only attempt at discriminating among different amounts of change in insight problems' representations has been that by Knoblich et al. (1999) on a categorical basis.

Insight problems' difficulty has been examined primarily in the context of specific problems, suggesting different problem-specific causes. Although the experience of insight is essentially seen as a "sudden restructuring," no other developments have been made to increase our understanding of this experience and what contributes to it – within and besides solving an insight problem. Scholars have not explored the effect of problem difficulty on the experience of insight.

The theoretical framework presented in the following section deals with some of the issues already identified. The framework of the intensity of the insight experience was developed by incorporating both structural and dynamic aspects of experience. Dynamic characteristics of insight problem solving were examined further applying Lewin's (1936) field theory, and a tentative measure of restructuring for the domain of matchstick arithmetic problems was developed. It is proposed that the difficulty of a problem affects the intensity of the insight experience. Further, it is argued that the difficulty of a problem depends not only on the structural properties of the problem itself, but also on the dynamic properties of the experience of solving the problem.

2. Theoretical framework

The concept of “insight” is treated in this work only with respect to problem solving. The term ***insight*** refers to “the process by which a problem solver suddenly moves from a state of not knowing how to solve a problem to a state of knowing how to solve it” (Mayer, 1995, p.3). The Gestalt psychologists referred to this process as “sudden restructuring,” emphasising the swift switch from one structure of the problem situation to a different one (e.g., Dunker, 1945; Scheerer, 1963/1967). A ***problem*** is defined using Dunker’s quote: “a problem arises when a living creature has a goal but does not know how this goal is to be reached” (Dunker, 1945, p.1). The term ***problem solving*** refers to the process by which an individual reaches the goal. These definitions are still used in the contemporary literature on insight (e.g. Gilhooly & Murphy, 2005, p. 279).

The concept of restructuring has already received a substantial amount of attention in the literature (e.g. Weisberg, 1995; Ohlsson, 1992; Ash & Wiley, 2004, etc.). The current work has taken the next step in developing the concept of restructuring by examining its effect on the intensity of the insight experience. Also, a measure of the degree of restructuring in the domain of matchstick problems is developed.

The term ***dynamic*** refers to forces that are acting on the individual and the resulting tension (Lewin, 1935, 1936, 1938). The dynamic aspect of the insight problem solving process has been largely ignored in the literature. Ormerod, MacGregor, and Chronicle (2002) have recognized both the importance of dynamics in problem solving as well as the fact that modern cognitive psychology has not paid sufficient attention to it. This thesis introduces the dynamic concept of ***tension*** developed by Kurt Lewin (1935) as an important aspect of the phenomenology of problem solving, which contributes to both the experienced difficulty of the problem and the intensity of the insight experience.

It is put forward in this thesis that insight is characterized by the intensity of the experience which depends on the degree of difficulty of the problem. The degree of difficulty is defined as the interdependence of the degree of restructuring involved in the problem and the amount of tension generated and released when the problem is solved. In this section, the effects of the degree of restructuring and the amount of tension on the intensity of the insight experience are analyzed and two hypotheses are formulated.

2.1. Degree of restructuring

The solution of an insight problem involves the formation of a new mental representation of the problem, a "restructuring" in Gestalt terms (Ormerod et al., 2002). This view is commonly accepted in the current research on insight (e.g. Davidson, 2003, Kaplan & Simon, 1990, Ohlsson, 1992, etc.): "Despite different emphases, the majority of approaches recognize that insight problem solving involves some kind of restructuring of the initial problem representation" (Chronicle et al., 2004, p. 14). The Gestalt notion of restructuring remains a key characteristic of this phenomenon, but there is still no consensus as to how the restructuring occurs and why it is difficult. The challenges in studying restructuring can partially be attributed to the wide variety of problems used from different domains that might have their own unique sources of difficulty of restructuring (Kershaw & Ohlsson, 2004).

The difficulty of insight problems lies in the fact that the problem solver forms an initially unproductive problem representation, which then has to be changed. Other things being equal, the bigger the difference between the initial representation and the goal representation (i.e. the bigger the restructuring required), the more difficult it will be to achieve. Knoblich et al. (1999) explained the difference in difficulty among several insight problems from the same domain by the difference in the amount of change required in the problem's representation. These authors distinguished among four types of matchstick equations based on the level of impact on the problem representation required for the solution. They hypothesized that problems that require more change to the initial problem representation would be harder to solve. Thus, Knoblich et al. (1999) introduced the idea of a positive relationship between the amount of change in the problem representation and the level of problem difficulty, supporting their hypotheses with experimental results.

In the following sections, the ***degree*** of restructuring involved in the solution of different problems in the same domain as used by Knoblich et al. (1999) is investigated further. The relationship between the degree of restructuring and the intensity of the insight experience is examined, hypothesizing that the degree of restructuring is one of the contributors to the problem's difficulty and the intensity of the insight experience.

> The **degree of restructuring** is defined as the amount of change to the initial representation of the problem required to transform it into the solution representation.

2.1.1. The amount of change in insight problems

The amount of change that a representation of a situation undergoes in one psychologically meaningful unit of action, or "step," can vary significantly in different situations. It could change slightly – for instance, by completing the sixth step in a ten-step task (such as washing ten dinner plates) – or, it could change significantly such as when a picture suddenly switches from being a picture of a young woman to a picture of an old lady in the classic example of an ambiguous figure (see Figure 1 for an example).

Figure 1: Classical example of an ambiguous figure: Young lady / old lady

For this picture to switch, it is enough to see only one particular element in a different way. For example, just seeing the necklace of the young lady as the mouth of the old lady momentarily leads to the change in the rest of the picture. This transformation of a single element (necklace to mouth) results in the transformation of the whole image. In one psychologically meaningful step (seeing necklace as a mouth), all elements of the picture assume a different meaning and organization, thus leading to a dramatic change of the representation.

In both of these examples - washing sixth plate out of ten and noticing that young lady's necklace could be seen as an old lady's mouth - a psychologically meaningful single unit of action resulted in the change of the representation of a situation. However, the amount of change in the representation that each of these steps caused varied significantly between the two examples. In the ten plate task example, an individual moved from being half-done to having only four left to do in one step. In the picture transformation example, in a single step an individual was looking at a completely different picture, and all elements of the picture had suddenly assumed different meanings and organization. This comparison illustrates the different amounts of restructuring the representation that can be achieved by one psychological step.

The discovery of the solution to an insight problem is experienced as happening suddenly. This suddenness implies that the necessary change from the current state to the goal state is achieved in one step of "sudden restructuring." That is, in insight problems, no gradual progress toward the goal is possible; the solver either knows the solution or does not know it, and there is no "in between" (e.g. Metcalfe& Wiebe, 1987). The solution to such problems is attained entirely at once, meaning that the change the problem representation undergoes from the initial state to the goal state is achieved in a single step. This concept is nicely captured in one-move problems such as matchstick equations.

With respect to the degree of restructuring involved, problems involving insight can be differentiated from other situations by two distinctive properties:

- All change in representation necessary to solve the problem happen as a result of one move. Even if the solution involves several steps, there is one "crucial" psychological step that leads to the necessary change of the representation;
- The amount of restructuring achieved over one move is significantly greater than in other situations (non-insight problems).

Knoblich et al. (1999) recognized that different insight problems involve different amounts of change to the problem representation in their solutions. For example, in the context of a matchstick equation, a change from the statement VII = V to the statement VI = VI is quite different than the change from the statement VII = I to the statement $\sqrt{I}$ = I; nevertheless, they both are caused by the relocation of the same stick leading to a balance in both statements.

Since the attainment of the goal in insight problems can be brought on by different degrees of restructuring the problem representation, the following hypothesis can be formulated.

Hypothesis 1: When two problems are solved under the same conditions, the intensity of the insight experience will be greater for the problem that requires the greater degree of restructuring.

While testing their constraint relaxation hypothesis, Knoblich et al. (1999) suggested that the greater the amount of change to the problem representation required by a solution, the higher the degree of difficulty of the problem. Thus, Knoblich and colleagues introduced a novel approach of examining the amount of restructuring involved in different problems. While experimental support for their hypothesis was reported, some issues could be raised with respect to the underlying assumptions of their quantification of the amount of change involved in different problems. To investigate the matter further, the suggested quantification by Knoblich et al. (1999) is now further discussed.

2.1.2. Quantification of restructuring proposed by Knoblich and colleagues (1999)

Knoblich et al. (1999) used a domain of matchstick equation puzzles in which numerous different problems could be generated. Matchstick equation problems are unbalanced mathematical expressions written using Roman numerals (e.g. V = III – III) that need to be brought to balance by moving only one stick (e.g. V = III + II). In their analysis of the domain of matchstick equations, the authors assumed a hierarchical three-level problem representation, and also suggested three constraints that might need to be relaxed to solve a problem. The authors proceeded to map the proposed constraints onto assumed levels of problem representation in order to identify which constraint would result in greater change to the problem representation.

2.1.2.1. Levels of problem representation

Analyzing their stimuli, the authors "assume(d) that the visual system parses matchsticks arithmetic problems into representation with three levels: numerals (I, II, III, etc.), functional terms (I + V; III - II, etc.), and entire equations (VI = V + I, III = II - I, etc.). The higher the level at which a change is introduced, the more encompassing is the resulting revision of the representation," (Knoblich et al., 1999, p.1537).

2.1.2.2. Constraints in matchstick equations

The authors also hypothesized that there are three constraints that could be operating in the domain of matchstick equations: value, operator, and tautology constraints. The value constraint refers to an assumption that a numerical value has to be changed only through operations, such as adding and subtracting the same quantity from both sides of an equation. The operator constraint refers to an assumption stating that arithmetic functions cannot be altered or deleted from an equation. Finally, the tautology constraint represents an assumption that an equation in a matchstick problem must have the form of Y = X + (or -) Z (assuming that tautological statements X = Z are meaningless).

Besides the constraints, another source of difficulty in the domain of matchstick problems according to Knoblich and colleagues could be the requirement to decompose a "tight chunk" - an element that has a greater degree of unity. For example, to decompose X or V was argued to be harder than III or II. A tight chunk is an element that is composed of sticks that, when decomposed, do not represent meaningful units in the domain of matchstick equations. According to this definition, only X and V elements were classified as tight chunks, but not "+" or "=" elements.

2.1.2.3. Mapping between constraints and levels of representation

The mapping was proposed between constraints and the levels of problem representation. It was hypothesized that the value constraint applies at the level of numerals, the operator constraint operates at the level of functional terms, and the tautology constraint applies at the level of the entire equation. Based on this mapping, three types of problems were distinguished with the tight chunk problems being the fourth type:

- type A problems that involve relaxation of the value constraint in their solution;
- type B problems that involve value and operator constraints;
- type C problems that involve operator and tautology constraint;
- type D problems that involve decomposition of a tight chunk, X ↔ V transformation.

Based on this classification, it was hypothesized that type A problems were easier to solve than type B problems, and type B problems were easier than type C problems. Also, based on the tight chunk hypothesis, type A problems were easier to solve than type D problems. However, the relationship of type D problems to problems of type B and C was neither specified nor tested.

2.1.2.4. Evaluation of the approach

As can be seen from the above description, the proposed differentiation of the degree of change among different equations is based on a categorization with some underlying assumptions that can be questioned. More specifically, the three levels of problem representation can be questioned; for example, an alternative explanation may be based on a figure-ground parsing with numerals processed as the figure level and the operations as the background level. Another assumption that can be debated is whether or not people impose the constraints suggested by the authors, since it is not clear how one identifies constraints operating in a given situation. Also, the mapping between the constraints and the levels of problem representation is not entirely specified. One can argue that a change of the operation in a matchstick equation operates on the level of the entire equation, but is only mapped onto the functional term level. Consider an example: The equation IV = III – I is solved into IV – III = I. In this case, according to the constraint relaxation hierarchy, only an operator constraint was relaxed and the problem was classified on the level of functional terms. However, it seems more reasonable to classify this change on the level of the *whole* equation, because the role of each of the numerals in the equation changed significantly as a result of a single move.

Knoblich et al. (1999) also hypothesized that once a constraint has been relaxed it will stay relaxed. However, it is not clear whether a relaxation of a higher level constraint (e.g. operator constraint) automatically leads to a relaxation of a lower level constraint (e.g. value constraint). If this is true, then the sequence in which the problems are solved by subjects has a significant effect on the results (solution rates and solution times). This issue was not discussed in the report of their study.

Given the concerns discussed above, it was decided not to use the taxonomy of matchstick equations as an independent measure of restructuring for the purpose of testing Hypothesis 1. Rather, a more subjective measure of perceived difficulty of a problem was used for this purpose.

Nevertheless, Knoblich and his colleagues have made a significant contribution. They introduced a domain of multiple comparable problems for the study of insight. Furthermore, they suggested a differentiation of problem difficulty based on the amount of change to the problem representation.

Building on the general idea of Knoblich et al. (1999) and using the same domain of problems, a more continuous measure of the amount of change was developed without grouping problems into categories. In the proposed measure of change, the initial state of the problem was compared to the final state of the problem. The degree of change was assessed based on the difference between the two representations. All the changes that were done to the initial representation in order to arrive at the final representation were identified and counted to compute the "change score" of a given problem with a given solution. Thus, a more gradual scale of measure of change was achieved while making fewer assumptions. The measure of change in matchstick equations is described in detail in the following section, and an initial, preliminary verification of the measure is reported in Experiment 1 of the Method section.

2.1.3. Matchstick equations and the measurement of change

Matchstick equations are unbalanced mathematical expressions that are represented with Roman numerals and basic mathematical operations (Knoblich et al., 1999). For example, a matchstick equation (IV – III = VI) corresponds to (4 – 3 = 6). A set of identical sticks is used to represent the equations. The goal of the puzzle is to balance the equation by moving only one stick. The stick can not be taken out from the equation, it must be placed somewhere inside the equation without doubling other sticks. At the end, the equation must be balanced and should represent a true mathematical statement.

Although a whole range of Roman numerals could potentially be used in matchstick equations, the allowable set was fixed to the range from 1 to 12. Possible operations included "+," "–,“ "×," "/," and "=." An "=" and at least two numbers, one on each side of the "=," must be present in each equation. Usually equations consist of three numbers and two operations, although other combinations are possible.

Matchstick equation puzzles allow generating a whole set of problems within the same domain. The problems in the set can vary significantly in the degree of restructuring required for the correct solution. The problems can be compared on the basis of the degree of change involved. The set used by Knoblich et al. (1999) was expanded to include operations such as “×” and “/” in addition to “+” and “–.“ In this domain, problems with multiple possible solutions could also be generated.

A procedure was developed for measuring the amount of change in a matchstick equation, which could be applied to any matchstick equation. This heuristic is based on the careful analysis of the stimulus. Using this procedure, one analyses what happens to the initial equation when a stick is taken away from one location and placed in another location. This analysis traces changes that happen to the elements and the relationships between them. This procedure permits constructing a more gradual scale of the amount of change than the categorization of problems into four groups proposed by Knoblich et al. (1999). The heuristic for measuring the amount of change in matchstick equations is described in detail in the following section which is preceded by a discussion on the relationships between the elements in the matchstick equations.

2.1.3.1. Relationships in matchstick equations

A matchstick equation can be seen as a system of meaningful elements that are interrelated. Each number or operation represents a complete and familiar unit to an observer and is considered to be a meaningful element in the equation. For example, “=” is seen as an equality sign and not as two separate sticks; “V” is seen as a number “5” rather than a letter “v” or a checkmark. The context in which the elements appear defines their meaning.

Relationships among the number-elements in matchstick equations are defined by the operation-elements in the equation. The relationship between two number-elements describes how change in one number-element affects the other number-element if all other factors are kept constant. This effect could be positive or negative. Two types of relationships between number-elements are distinguished and labeled as positive relationship and negative relationship.

A ***positive relationship*** between two number-elements in a matchstick equation is defined as a relationship when change in one of the elements in a certain direction leads to a change in the same direction in the other number-element, while everything else is kept constant. If two elements, *A* and *B* are in a positive relationship, then, if *A* increases, *B* also increases; if *A* decreases, *B* also decreases.

A ***negative relationship*** between two number-elements in a matchstick equation is defined as a relationship when change in one of the elements in a given direction leads to a change in the opposite direction in the other number-element, while everything else is kept constant. If two elements, *A* and *B* are in a negative relationship, then, if *A* increases, *B* decreases; if *A* decreases, *B* increases.

For example, in the equation IV + II = VI the relationship between the elements "4" and "6" is positive since if element "4" increases by three to become "7," the element "6" will also increase by the same amount (becoming "9") if all other elements are kept constant i.e. number-element "2" and both operation-elements are not changed. Following the same logic, the relationship between the elements "2" and "6" is also positive. However, the relationship between the elements "4" and "2" is negative since increasing the element "4" will lead to a decrease in the element "2" to keep the whole expression valid (i.e. the element "6" and both operations are kept unchanged).

Both positive and negative relationships could also be of different magnitudes. The magnitude of a relationship depends on the operations involved. Operations "+" and "–" lead to exactly the same amount of change in both elements involved in a relationship. In the above example, if element "4" is increased by 3, then, the element "6" will also increase by 3, leading to exactly the same amount of change in both numbers. However, operations "×" and

"/" (this will also apply to a square root operation if the set is extended to include it) result in different amounts of change in the two participating number-elements. For example, in the equation $4 \times 2 = 8$, an increase in the element "4" by only 1 unit will result in an increase in the element "8" by 2 units, thus, leading to unequal amounts of change in the two number-elements.

Two levels of magnitude in the relationships between number-elements are distinguished and labeled as equal and unequal. Under an ***equal*** effect, a certain amount of change in one element causes exactly the same amount of change in the other element. Under an ***unequal*** effect change by a certain amount in one element causes a different amount of change in the other number-element.

The type of the relationship between the number-elements and its magnitude can only be changed by changing or adding an operation.

There are also other possible relationships in an equation, e.g. relationships between a number-element and an operation-element or between two operation-elements. For example, in an equation $4 + 2 = 6$, the possible relationships could also include a relationship between the number-element "4" and the operation-element "+,"or a relationship between two operation-elements "+" and "=." A change in a number-element does not affect an operation-element at all. A change in operation does not directly affect any given single number-element; nevertheless, it influences the relationship between a pair of number-elements. The effect of the operation-element change on the number-elements is captured in the change in relationships between pairs of number-elements in an equation. The relationships between two operation-elements (e.g. relationship between "+" and "=") are undetermined and are not considered here since it is not clear how change in one operation affects other operations in an equation.

2.1.3.2. Measuring change in a matchstick equation

Change in a matchstick equation is measured by comparing the solution representation of the problem to its initial state and counting all the changes that took place in this transition. By carefully analyzing what happens when a stick is removed from one location and placed into a new location in the equation. The changes of the elements and relationships between them are tracked and the number of changes is counted. The total number of changes corresponds to a problem's ***change score***. The same heuristic could be applied to every problem in a set of matchstick equations, and the scores for different matchstick problems could be compared.

Removing a stick from one location in the equation and placing it in another location results in a series of changes rather than a single instance of change. Besides changing the element from which the stick was removed and the element where it was placed, a series of other transformations take place. Some transformations may be more difficult than others (e.g. decomposing a "V" or conceiving of a second equality), which is reflected in the number of units assigned for different transformations. Table 1 presents the heuristics used in assigning units of change for each possible transformation and the rationale for each change score.

In the analysis of a particular equation, all changes are identified and counted. For example, equation (IV – III = VI) is balanced by taking a stick away from the number 3 and replacing it on top of the minus sign, thus, the solution is (IV + II = VI). The list of all the transformations that take place when the initial state of this equation is changed into its solution state is presented below.

Table 1: Heuristics and rationale for counting the change in matchstick equations

Action	Units of change	Rationale
Removing vertical "I" from a number	1	When a stick is removed from a number, that number changes
Decomposing "V," "X," "+," or "="	2	These elements represent a more cohesive unit, which is harder to decompose than to take away a vertical stick from a number, e.g. "II." Assigning a higher value to such decomposition was inspired by a "tight chunk" idea of Knoblich et al. (1999)
Adding a vertical "I" to an existing number	1	When a stick is added to an existing number, that number changes
Recombining the sticks to make "V," "X," "+," or "="	2	When a more cohesive unit is constructed, more change happens as opposed to when a vertical "I" is added to an existing number
Changing a part of a number into a part of an operation or vice versa	2	E.g. A vertical stick from "IV" becomes a part of a "+." In such a transformation one has to stop seeing "I" as a number (one change) and then change the meaning of the stick into a part of an operation (another change)
Changing a part of one operation into a part of a different operation	1	The meaning of an element has to change, e.g. "-" becomes a part of "=." have to change the meaning of "-" as a minus into a part of the "="
Starting the transformation from an operation element	1	In the context of a matchstick equation, value elements are more likely to be seen as the figure and operations are put into the background. Starting the transformation from the element from the background might be more difficult than starting the transformation from the element in the figure
Changing the overall value in the equation	1	If a transformation results in a change in the overall value in the equation it constitutes a change (e.g. [VII - II] changes into [VI - III] versus VI changes into IV).
Changing the relationship between two elements	1	If a relationship between two elements changes from positive to a negative or vice versa, or if the magnitude of the relationship changes, it constitutes a change.
Changing the existing "="	2	"=" is often seen as a "fixed" element in an equation, which has to remain unchanged to satisfy the balancing condition.
Creating a second "=" in an equation	2	This changes the equation into a form of A=B=C and therefore changes its structure considerably
Creating a new element in the equation	2	When a new number-element or a new operation-element is added to the equation (as opposed to changing the value of an existing number-element or meaning of an existing operation-element) it constitutes a change in the representation

- The vertical “I” is removed from number “III” (1 unit);
- The vertical “I” is changed from being a part of a number into being a part of an operation (2 units);
- “–” changes from being a minus sign to a part of a “+” sign (1 unit)
- Two sticks are combined into a “+” (2 units);
- The relationship between “IV” and “III” changes from positive into negative; the relationship between “III” and “VI” changes from negative into a positive (2 units);
- The total value in the equation is changed (1 unit)

Total change score: 9 units.

Clearly, the change score computed through this procedure is an approximation to the actual structural change of the problem. This heuristic allows for ranking of the matchstick problems with respect to the amount of structural change required to balance each equation.

The computation of the change score results in a more continuous scale of measurement of the degree of restructuring in matchstick equations than the typology proposed by Knoblich et al. (1999). Four potential benefits provided by the continuity of the scale are:

- It allows differentiating between problems that were classified as the same type by Knoblich et al. (1999), thus, predicting differences between the problems in the same category;
- It allows for making comparisons between problems of “tight chunk” type to problems of “operations” type and to the problems of “tautology” type. These comparisons are not possible based on the categorical distinctions;
- It allows for extending the set of matchstick problems (i.e. including other operations or other forms of equations) and making comparisons between them.
- Change scores can be adjusted based on empirical evidence, thus increasing the correspondence between the problem difficulty and the measure of restructuring.

In the development of the change score measure, the analysis was done as carefully as possible; however, the proposed heuristics still might require certain adjustments. The initial set of data for the preliminary verification of (and possible adjustment of) the developed measure of restructuring was collected as part of Experiment 1.

However, the degree of restructuring required in a problem is not the only factor affecting the *experienced* difficulty of the solution and the intensity of the insight experience upon finding the solution. The same problem could become easier or more difficult depending on the environment where it is solved, the importance of solving this particular problem to the individual, and other such factors. The same degree of restructuring can be achieved with different degrees of difficulty by different individuals. The same amount of restructuring can be made easier or harder by modifications such as varying the amount of time available to work on the problem, changing the relative importance of solving the problem, or providing different information about the problem. While these factors do not affect the degree of restructuring involved in the solution of the problem, they significantly affect the dynamics of the solution process and the experienced difficulty of the problem. Consequently, the difficulty of a problem and the experience of insight depend on both the restructuring required to solve the problem and the psychological conditions governing the process of finding the solution.

Put another way, the insight problem solving can be seen as a process of working towards a goal that has certain resistance on its path. The path resistance depends not only on the structural properties of the problem, but also on the psychological environment an individual finds themselves in during this process. How one experiences the discovery of the solution, depends on the aggregate resistance one encounters on the path towards that solution.

Therefore, to more completely understand what determines the intensity of the insight experience, it is worthwhile to investigate the immediate psychological situation of a person both while they are seeking a solution and at the moment when they find it.

2.2. Dynamic component of the insight problem solving process

The immediate psychological situation at the moment of solution depends heavily on what happens while the problem is being solved; that is, the process of struggling and finding the solution. It is generally accepted in the field that the state of impasse is one of the distinguishing characteristics of the insight problem solving activity (e.g. Ohlsson, 1992; Knoblich et al., 1999, etc.), but is not a necessary condition of insight (e.g. Ormerod et al. 2002). In their analysis of the modern theorising of insight, Ormerod et al. (2002) point out many similarities between contemporary explanations of insight and the original Gestalt account, with one significant difference. This difference, according to these authors, is the lack of attention in modern literature to the dynamics of the process of solving an insight problem. Ormerod et al. (2002) contrasted the current view of impasse as a “quiescent or inert” state to the state of conflict under high tension (Lewin, 1935, 1936), suggesting that this latter view “may have been closer to the truth” (p.792).

The present theoretical framework incorporates an aspect of insight problem solving that has been missing in the modern theorizing of insight: The presence of a state of conflict under high tension (Ormerod et al. 2002). In the following section, the dynamic property of tension in the insight problem solving process is analysed applying Kurt Lewin’s field theory. A method of studying the effect of tension on the intensity of the insight experience is described in the Method section, Experiments 3 and 4.

2.2.1. Field theory of Kurt Lewin

Kurt Lewin is known as one of the prominent social psychologists, but his study of individual cognitive processes is often overlooked despite his significant contributions in the study of individual behaviour (Hall & Lindzey, 1978). Lewin developed field theory, which is a body of definitions and axioms, and presents a system of interdependent psychological concepts that are equally applicable to a variety of fields in psychology (Lewin, 1935, 1936, 1938). Lewin introduced such dynamic concepts into psychology as force, tension, valence and barrier. Lewin did not regard his field theory to be a “theory” in the conventional sense;

according to him the field theory could not be right or wrong. Lewin wrote: "*Field theory is probably best characterized as a method*: namely, a method *of analyzing casual relations and of building scientific constructs*" (Lewin, 1951, p. 45, emphasis original).

The application of field theory to the study of behaviour led to a novel investigation of human activity. Some examples of the hypotheses generated from Lewin's theory that were experimentally tested include the remembering of unfinished tasks (Zeigarnik effect), resumption of unfinished activities, and different reactions to frustration (Gold, 1992). The cognitive dissonance theory of Leon Festinger was predominantly built on the notions of filed theory without explicitly using its language. Field theory allows one to provide a theoretical explanation of various aspects of human behaviour through its system of interdependent concepts.

However, no theory can avoid being criticized, and field theory is no exception. The adoption of terminology from physics and mathematics created confusion among many even though Lewin had always defined and used the terms in a purely psychological sense and had warned against premature formalization (Back, 1992). Field theory was also criticised for neglecting to take the past history of an individual into account when predicting behaviour, since the theory emphasized the immediate situation of the individual. Lewin stressed that field theory was misunderstood in this respect (Lewin, 1951; Hall & Lindzey, 1978). According to Lewin, the totality of the immediate situation of a person is the sole determinant of behaviour at a given time, but the immediate situation includes relevant elements of the past to the extent of their effect on the present situation (e.g. present tension systems, force fields, needs and valences) (Lewin, 1951).

Lewin did realise that his theoretical framework was just a beginning and that concepts of field theory "will certainly have to be revised in the course of time" (p. 7, Lewin, 1936). In spite of its advantageous explanatory power, field theory has not been adopted in the context of studying individual psychology. This can be partially attributed to the fact that Lewin himself turned his attention to studying social aspects of behaviour (Gold, 1990). For example, Hall & Lindzey (1978) observed that "there have been no important advances in

Lewin's theory of the person since the early 1940s" (p. 425). The fact that the major constructs of field theory, such as forces and valences, were not directly observable – indeed, one could only witness their effects – posed difficulties in measuring and quantifying such aspects. This contributed to a general aversion towards using the terminology associated with field theory and eventually towards the theory altogether. Psychology turned to more objective aspects of behaviour. The information processing approach became the major trend in studying cognitive phenomena, offering the convenience of computer simulated human thinking that did not have pose the same difficulties as measurement of forces.

This has contributed to the fact that the dynamics of the process and the experience of the individual are missing in the current theorising of cognitive phenomena and insight problem solving in particular (Wieth & Burns, 2000). Ormerod et al. (2002) suggested that Lewin's conceptualisation of the state of impasse "may have been closer to the truth." Even though Lewin himself did not analyse the phenomenon of insight in detail, many of the concepts developed in his theory are highly applicable to this process. It is worthwhile to revisit Lewin's conceptualization of behaviour of an individual applying it to the insight problem solving situation. Field theory provides the necessary constructs to analyse the dynamics of the process of solving a problem and to formulate a representation of the immediate psychological situation of the individual.

2.2.2. Immediate psychological situation

Kurt Lewin theorized that person's immediate psychological situation, which incorporates both the individual and the environment, determines the behaviour. The immediate situation is represented as a *life space* with different *regions of activity* that psychologically exist for the individual. A life space is the dynamic and constantly changing *psychological environment* of an individual. At any moment in time, a person is located in one of the regions of activity. In other words, an individual is always doing something. It could be walking to the library, looking out the window, thinking about a problem or having an argument. A region of activity refers to a *psychological* activity rather than *physical* activity although at times the two might coincide. Regions of activity can be further differentiated into sub-regions that constitute psychologically meaningful units of action. *Locomotion* refers to the movement within the life space from one region of activity to another, or between sub-regions of a given region. Change within the life space is also considered locomotion.

Regions of activity are characterised by their *valence*, or their attractiveness to the individual. Valence is a dynamic property of a region that can be positive or negative. Regions in the life space acquire their valences based on needs, physiological states, wishes or intentions. Valences create *force fields* in the life space that steer the locomotion through the life space. A positive valence creates a force field that attracts to that region, while a negative valence creates force field that pushes away from that region. The strength of a force changes depending on the distance to the region causing it. When the valence of a given region stays intact, increasing the distance from that region decreases the force associated with it. Forces away from a negative region decrease at a faster rate than forces towards a positive region. A variety of different forces exist within the life space at any given point in time. The locomotion through the life space is determined by a totality of all the forces acting on the individual at that moment in time, called a *resultant* force.

Tension is another property of a region of activity. Tension arises in a life space as soon as there is an unsatisfied need or a wish. There is a tendency towards an immediate discharge of tension as soon as it arises, which gives energy to and sustains all mental activity of an individual. The tendency to seek a discharge of tension as soon as it arises steers the behaviour to satisfy the need causing it. As soon as the need is satisfied, tension is discharged. The extent of tension is determined not only by the extent of the need producing it but also by the friction of opposing forces in the field. This means that greater resistance on the path to the goal causes a greater state of tension in the system.

The resistance on the path can be caused by different *barriers* that could be either physical or psychological. For example, a fence is a physical barrier that hinders locomotion through it, while an adult's prohibition to eat the chocolate that is on the table is a psychological barrier that deters a child from reaching their goal of eating the chocolate. The possibility of locomotion towards the goal might be prevented upon reaching a barrier. This would lead to a temporary state of *conflict.* The state of conflict is characterized by an inability to move, since the resultant force in this state is equal to zero. This state is also characterized by an increasing tension as the opposition of the forces is at its highest level. Overcoming a barrier might involve getting through the barrier with more effort, if the barrier is passable, or it might involve going around the barrier, finding a *detour*, if the barrier is impassable. Finding a detour requires overcoming the acting forces and a restructuring of the life space in such a way that the whole path to the goal becomes present in the field.

A more detailed explanation and summary of the field theory concepts that are relevant to the present work can be found in Appendix A. For a more elaborate discussion on this theory please refer to the source, Lewin (1935, 1936, 1938, and 1951).

2.2.3. Dynamics of the process of solving an insight problem

Lewin (1935) viewed the phenomenon of insight, central to the present work, as an abrupt structural change of the life space of an individual that happens in the context of a detour. In the act of insight, “the structure of the field as regards its grouping into wholes undergoes a transformation, usually an abrupt one… the act of insight…consists in a transformation of the whole relations in the field” (Lewin, 1935, p. 196). The notion of restructuring, discussed in the previous section and emphasised by many scholars of insight, in Lewin’s terminology refers to the structural change of the region of the problem in one’s life space. However, Lewin himself, concerned with numerous other aspects of human behaviour, did not analyse the phenomenon of insight in greater detail. Insight, as with every other psychological event according to Lewin, happens in the context of, and ultimately as a result of, the state of the life space at that moment.

The dynamics of the process of solving an insight problem, and what happens at the moment of solution in particular, is analysed below in the context of an example from the domain of matchstick arithmetic problems (Knoblich et al., 1999). The purpose of the following discussion is to investigate the dynamics of solving an insight problem, especially the role of tension in this process.

2.2.4. Dynamics of tension in insight problem solving

The insight problem solving process can be conceptualized as locomotion through a region of activity towards the goal. Such locomotion is characterised by the exploration of an initially unstructured region of the problem's search space. The exploration of an unstructured region in the search for a path to the goal potentially leads to entering sub-regions that are "dead ends." This results in encountering a barrier, or a series of barriers, and arriving at a temporary state of conflict (Lewin, 1935) in the field at some point during this process. The solution of an insight problem has detour properties and requires the restructuring of the field.

The events of entering the problem solving situation, encountering a barrier, reaching a state of conflict, and restructuring the field to achieve a successful detour can all be viewed as processes that change the state of *tension* in the system.

> **Tension** is defined as psychological pressure produced by a need or a wish to act in order to satisfy that need or wish.

2.2.4.1. Structure of the life space at the beginning of the problem solving process

A problem solving situation can be represented as a region of activity in a person's life space. There can be many possible reasons for a person to enter the problem solving region. For example, a person might enter the problem solving region because of participation in an experiment on problem solving. In this situation, other factors aside, a socially induced force will exist for the person to remain in the problem solving region. Also, the fact that the person is being observed might potentially affect the valance of solving the problem. These issues might not be present if the problem were attacked in a privacy of one's home without an observer. The reasons for entering a problem solving region have an influence not only on the structure of this region of activity in the life space, but also on the dynamics of the problem solving process. These reasons could determine, at least to a certain degree, the level of tension in the system, the strength of the valance of the goal region, the firmness of the boundary of the problem solving region, and its structure as well.

Initially, the problem solving region (PSR) exists for a person as one unstructured cell. Before the person enters the problem solving region, it is not divided into sub-parts. For example, an individual that participates in an experiment on problem solving has a region of activity, "participation in the experiment," in their life space. Upon knowing that in the experiment they will be asked to solve a problem, the "participation in the experiment" region of activity has a sub-region, "solving a problem." However, when the individual had just arrived to the experimental lab, they did not have much information about the problems they have to solve. The structure of the "participation in the experiment" region of activity might have consisted of the following general sub-regions: "come to the lab," "get some information about the problem and ask questions," "solve the problem," "get the bonus mark," and the end.

As the person enters the problem solving activity itself, the "problem solving region" becomes more differentiated, and is broken down into sub-regions. The PSR acquires a sub-region of the initial state of the problem, and it is through this that the person enters the PSR. For example, a matchstick equation IV = III – I (Knoblich et al., 1999) needs to be balanced by moving only one stick. When the problem solver learns that this is the problem they need to solve, the initial state of the equation, IV = III – I, becomes the sub-region in the PSR, which is further decomposed into sub-parts (this is discussed later). At the same time, the PSR acquires a goal sub-region which corresponds to the solution of the problem. In the above example, the goal sub-region is not well defined. The equation has to be balanced, which implies that the right side of the equality should be equal to the left side, but the exact form of the equation and specific numbers in the equation are not specified. Thus, the goal sub-region is not well structured.

Along with the goal sub-region, the PSR might acquire a "Didn't solve" sub-region that corresponds to the failure to find the solution to the problem. The existence and salience of the "Failure" sub-region greatly depend on the problem solving environment. The existence of the "Failure" sub-region depends first of all on the presence of a point in time when the problem solving activity will be terminated. At that point, the problem can either be solved or not solved, and the individual will have to stop working on the problem. The psychological consequences associated with not solving the problem further determine the characteristics of the "Failure" sub-region. For example, a student has to solve a given problem in the context of a one-hour exam. From the beginning of the examination, the possibility exists for the student that they will not be able to solve the problem within the given examination time. That possibility is undesirable as it will result in a lower course mark. Furthermore, if the student's scholarship depends on this course grade, the valence of the "Failure" region will become even more negative. The existence of the failure sub-region, its proximity to the person and its valence are greatly determined by the conditions and the environment under which the individual enters the PSR. For example, being in an experimental situation and having a limited time to work on a problem might make the failure region more salient than it would have been if the problem were solved in the comfort of one's home. The valence of this region will be negative and its strength will increase with the importance of finding the solution.

The sub-region of the initial state itself assumes a structure that corresponds to the grouping of the elements of the problem and their relationships based on their psychological existence for the person at that time. In the context of the equation presented above, the sticks used in the equation are not seen as a mere collection of 10 sticks, but rather as sticks grouped into meaningful units (numbers and operations) that form the equation. The sign "IV" is one such unit in the equation, representing the number 4. The three sticks forming sign "IV" are seen as parts of this sign, and all three of them form the number 4. The same applies to sign "III", which stands for the number 3 in this equation and is composed of three sticks. The element "I" is seen as 1; two sticks forming an equality sign are seen as such "="; and the stick "–" represents a minus sign. Therefore, the whole equation is seen as 4 = 3 – 1. The elements in this equation are also engaged in certain relationships among each other. For example, 4 is the outcome of the subtraction operation on the right side. Figure 2 provides a possible representation of the structure of the problem solving region at the very beginning of the problem solving process.

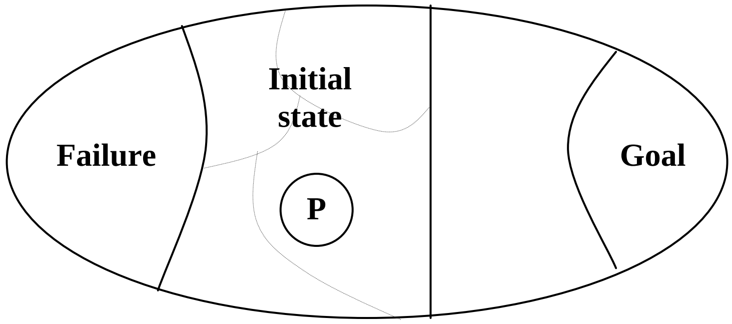


Figure 2: Structure of a problem solving region with a Goal, Failure, and Initial State sub-regions, the person (P) is located in the initial state

The initial state of virtually any problem situation is, by definition, an unbalanced state. This means that any problem is formulated such that one wants to alter or move away from the initial state of the problem to achieve the goal. The initial state of a problem is an “undesirable” state and the solution of the problem is a “desirable” state. The initial state of the problem sub-region, having a negative valence, produces a force away from itself, but not necessarily towards the goal. The goal sub-region in the PSR has a positive valence and creates a psychological force towards it (represented as $F_{G(+)}$ in Figure 3). The strength of the valence of the goal sub-region, and of the force created by it, depends on the importance the person puts on solving the problem. The greater the valence, the stronger the force would be. The positive valence of the goal sub-region creates a unidirectional force in the life space that pushes towards that sub-region.

As a result, a state of tension exists in the PSR, meaning that there exists a need to change or move away from the initial state of the problem to reach the goal. This tension can naturally be relieved by relocating to the goal sub-region. However, in a problem solving situation, the sub-region of the initial state and the sub-region of the goal state are not connected in the life space, and are separated by an unstructured space (Figure 3). The unstructured space between the goal and the present sub-region makes the immediate locomotion between these two sub-regions impossible, and the goal unattainable at first.

The failure sub-region in the life space has a negative valence and creates a force field that pushes away from that sub-region (represented by many outward arrows in Figure 3). Force produced by a negative valence does not have a specific direction. Instead, it operates in all possible directions that are away from that region. The strength of the valence of the failure sub-region depends partially on the consequences of not solving the problem that the individual paints for themselves, and partially on the strength of the positive valence of the goal sub-region. In some cases, the failure sub-region could be a negative mirror reflection of the positive goal sub-region. In other cases, the failure region might not even exist for the individual in the beginning of the process and might appear in the life space at a later stage, or not appear at all. Yet in another case, the negative failure sub-region could have valence far greater than that of the goal sub-region.

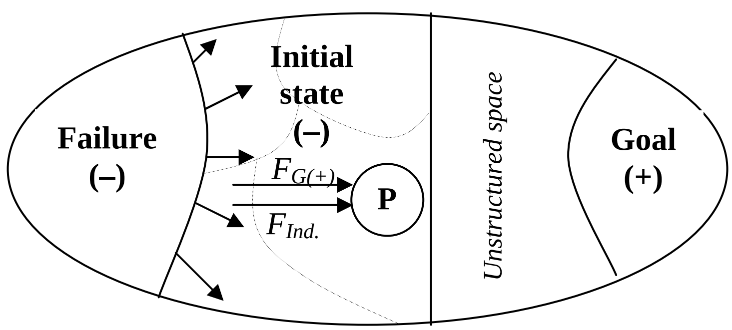


Figure 3: Problem solving region at the beginning of the problem solving process

For example, consider a situation when students are invited to participate in an experiment on problem solving in exchange for bonus marks towards their course grade. When a student is solving a problem in the context of this experiment, they might not want to "look bad" in front of the experimenter failing to solve the problem. This will create the negative failure sub-region in their life space, and the valence of that region will depend on the significance they place on "not looking bad in front of the experimenter." Moreover, if the student also thinks that they will not get their bonus marks if they do not solve the problem, then the negative valence of the failure sub-region will increase even more. On the other hand, if the student encountered the same problem in a magazine instead of the experimental situation, and decided to tackle it "just for fun," the failure sub-region might not have even exist in their life space.

Furthermore, in the experimental situation, there also exists a socially induced force acting on the student to solve the problem (represented as $F_{Ind.}$ in Figure 3). The experimenter selects the problem, presents it to the student and expects the student to solve it. The expectation on the part of the experimenter constitutes a socially induced force in this situation. If the student encountered the problem in a magazine and were solving it "just for fun" out of their free will, then there would be no socially induced force.

In summary, there are several force fields that might operate in the problem solving region at the beginning of the problem solving process. These are:

- The unidirectional goal force produced by a positive valence of the goal sub-region;
- The force away from the initial state sub-region produced by a negative valence of the initial state of the problem;
- The socially induced force coming from the environment;
- The force away from the failure sub-region produced by a negative valence of the failure sub-region.

The behaviour of an individual is steered by the resultant of all the forces present in the field at the time (Lewin, 1935, 1936, 1938). In the problem solving situation described above, the resultant force will be determined by the interaction of the four force-fields. The resultant force will be in the direction towards the goal, mainly due to the presence of the unidirectional goal-force and the induced force which operate in the same direction. The power of the resultant force will depend on the power of these two forces, and will be strengthen by the forces away from the failure and the initial state sub-regions, as these forces are aligned with the direction towards the goal.

2.2.4.2. Dynamics of the process of looking for a solution

Finding themself in an unbalanced state with increased tension, the individual starts acting to relieve the tension. The tension in a problem solving situation can be naturally relieved by reaching the goal and solving the problem. Therefore, the individual starts moving in the direction that they consider being towards the positive goal sub-region and away from the negative failure sub-region. Driven by the resultant force in the direction of the goal, the problem solver starts exploring the unstructured psychological space that lies between them and the goal. This exploration is not random for the individual; the driving force leads behaviour toward path that appears to most directly lead to the solution in the current structure of the life space (Lewin, 1936; Ormerod et al., 2002). According to Lewin, one of the distinct properties of insight problems is the requirement of a detour. That is, the initial, seemingly the most direct path toward the goal usually does not lead to the solution but to a dead-end.

The present structure of the initial state sub-region defines the possible set of sub-regions adjacent to it. These adjacent sub-regions are the regions of possible modifications of the initial state, i.e. available actions. The possible locomotion within the PSR is limited to this set of adjacent regions. The adjacent regions in the example of the matchstick equation discussed above could be "decrease 4," "increase 3," change "–" into "+," and so forth. In other words, the moves that the person is able to try depends on the structure of the initial representation of the problem sub-region.

The problem solver moves into one of the adjacent sub-regions in an attempt to reach the goal. For example, they first might try reducing the element "IV" by removing a vertical stick; let us call this an 'attempt 1' sub-region. However, as the vertical stick is lifted, the individual realizes that the remaining sticks actually form a number that is even greater than 4 – the Roman V representing 5. Moreover, replacing this stick to other locations, such as adding it to "I" or putting it over "–" to make it into "+," does not make the two sides equal.

As the 'attempt 1' fails, the person finds themselves in a "dead-end" and encounters an impassable barrier that blocks the goal. Encountering an impassable barrier increases the tension in the system. The barrier is a source of restraining force that is exerted on the individual upon encountering it. The restraining force of the impassable barrier is equal to the resultant force toward the goal, and thus the tension increases. The magnitude of the tension increase depends partially on the strength of the forces in the field and partially on the structure of the region and the availability of other “possible” moves to try that might potentially lead to the solution. Along with the increasing tension, the valence of the goal sub-region might also increase if the person views the problem to now be more challenging and applies greater effort to solve it. If this happens, then the resultant force towards the goal increases and, therefore, the restraining force of the next potential barrier increases as well.

In addition to the change in the state of tension and the possible change in the valence of the goal sub-region, the problem solving region also undergoes a structural change. The sub-region of ‘attempt 1’ was initially seen as being in the direction of the goal and having a positive valence. However, after realizing that the goal can not be reached through the ‘attempt 1’ sub-region, it is now seen to be in the direction away from the goal and towards the failure sub-region. In other words, the ‘attempt 1’ valence shifted from positive to either neutral or negative. Depending on the level of tension in the region, the ‘attempt 1’ sub-region might even become a part of the failure sub-region. This is more likely to happen when the tension in the region is relatively high, and the boundaries within the region are more easily destroyed.

Besides the repositioning of the ‘attempt 1’ sub-region to be in the direction away from the goal, other changes in the life space are also possible. Any repositioning of a stick leads to a certain amount of change to the initial representation of the problem. For example, removing the vertical stick from “IV” and replacing it next to “I” changes the values of both “IV” and “I” while other items and relationships in the equation remain intact. If the vertical stick from “IV” were put over “–” sign making the “–” into “+,” then “IV” would change its value, the sign would become its opposite, and the new element “V” would become a result of a summation rather than a subtraction operation. Changing the operation in the initial equation might give the individual the idea of trying other possible alterations to the sign, such as changing “–” into “/,” thus creating another sub-region in the direction towards the goal. These two examples illustrate that one move of a stick might result in different changes, or restructurings, in the representation of sub-regions of the problem. These changes define the possible set of adjacent sub-regions, which constitute the realm of available moves.

When the initial attempt fails, the person moves to another sub-region adjacent to their current position, and tries reaching the goal from there. For example, one might try to "increase 3" or change "–" into "+," both of which are also doomed to fail. If the second attempt is unsuccessful, a barrier is reached again, and the tension is similarly increased. If the valence of the goal had increased after the first attempt, then the tension increases to an even greater extent. Again, the magnitude of the tension increase depends on the strength of the opposing forces and the structure of the region.

As the attempts fail, more and more sub-regions are repositioned in the direction away from the goal and become potentially combined with the failure sub-region. When all possible sub-regions adjacent to the initial state are exhausted or seen as blocked, and no new ones appear, the individual finds themselves again being separated from the goal by an unstructured space with no possibility of locomotion towards the goal. The unstructured space now acts as an impassable barrier on the path to the goal (Lewin, 1951, p. 255).

The lack of available sub-regions in the direction towards the goal leads to a state of impasse or conflict. The state of conflict is characterised by increased tension due to the opposing forces and by the inability to move within the problem solving region (Lewin, 1935). The resultant force in such situations is zero since all the forces in the field cancel each other out. The nature of this conflict comes from the fact that the individual is blocked from reaching the desirable goal sub-region. The individual is trapped between two forces, one pushing toward the goal and the other pushing away from it with the same strength, and no movement in the direction of the goal seems possible at that moment. This state of conflict is a state of a "quasi-stationary" equilibrium under high tension (Lewin, 1935). If an impasse is understood to mean a temporary inability to move within the PSR region, then a series of impasses are possible (but not necessary) during an insight problem solving experience.

The barrier, by itself, is not negative at first, but with repeated encounters it acquires a negative valence that increases with every unsuccessful attempt (Lewin, 1935). The barrier's negative valence generates a force in the direction away from it. The presence of this force away from the barrier in the field increases the level of tension and eventually leads to a decrease in the magnitude of the resultant force in the direction toward the goal. The person will continue trying to reach the goal until the resultant force is still in the direction toward the goal through the barrier.

The evidence of increased tension in the system due to reaching a state of impasse can be found in the results of the Seifert et al. (1995) study. The authors replicated the Zeigarnik effect in the domain of problem solving, reporting the highest recall rate for the group of subjects that was interrupted after reaching a state of impasse. According to Zeigarnik, unfinished tasks are better remembered because tension was not discharged in those regions of activity due to interruption. The greater the unrelieved tension, the better the memory of the activity. Upon reaching an impasse, tension increased to a higher level than before reaching this state, and thus led to superior recall in the Seifert et al. (1995) study.

According to Lewin, increasing tension in one of the sub-regions has a tendency to spread to neighbouring sub-regions if the boundaries permit for the communication of tension. The boundaries among different sub-regions in the PSR have a higher degree of communication than between the PSR itself and other regions in the life space. The increased tension in one of the sub-regions in the PSR will very likely spread to other sub-regions within the PSR, depending on the level of tension and the firmness of the boundaries. The spreading of tension might result in the "flooding" of the region with tension and the wiping out of boundaries among the different sub-regions within the PSR. If the boundaries between the different sub-regions in the PSR are destroyed, the orderly search for a solution becomes even more difficult. The differences between moves that were tried or thought of disappear and everything becomes one undifferentiated mass. Increasing tension might affect the structure of the life space adding to the experience of difficulty of solving the problem.

The difficulty of finding the solution to insight problems depends on the strength of the acting forces in the field (Lewin, 1951). This is mainly due to the detour property of insight problems. On one hand, the cognitive restructuring required by an insight problem becomes more difficult because the person has to ***move against*** stronger forces. On the other hand, stronger forces result in a state of higher tension. After a certain level, higher tension might lead to "a narrowing-down of the psychologically existing area" and a possible "primitivation (regression)" which makes the restructuring more difficult (Lewin, 1951, p.254).

The evidence of the detrimental effect of the increasing forces in the field on insight problem solving could be found in the studies of Glucksberg (1962) and McGraw and McCullers (1979). Both authors reported an increase in solution time of insight problems for participants who were offered an incentive for their performance. Added incentive for performance in these studies could be viewed as an increased force in the direction towards the goal which contributed to the problems' difficulty through an increased level of tension.

One of the main principles of field theory is that the behaviour is steered in the direction of tension reduction. An increasing tension in the life space creates a pressing need for the individual to act to relieve this tension. In a problem solving situation, the release of tension is sought through reaching the solution to the problem. When a conflict state, as described above, arises during problem solving, the attainment of the solution seems to be blocked for some time, preventing a "natural" releave of tension in the region through the attainment of the goal. In line with field theory, the action towards the tension reduction will nevertheless persist. If achieving the goal does not seem possible, the tension reduction action might take the form of adjusting one's cognition[1] in the direction of tension reduction as outlined by the cognitive dissonance theory (Festinger, 1957).

Being in a state of increasing tension, the person might attempt to reduce the tension in the region, or at least to slow down the rate of its increase, without reaching the goal. This

---

[1] The term ***cognition*** is used as defined by Festinger (1957) referring to "any knowledge, opinion, or belief about the environment, about oneself, or about one's behavior" (p. 3)

could be achieved by re-evaluating the valence of the goal and/or failure sub-regions. By reducing the valence of reaching the goal or presenting the failure to be less undesirable, the strength of the forces in the field will decrease, thus stabilising or even potentially decreasing the overall state of tension in the PSR. Stabilizing the state of tension in the PSR will ease the task of carrying out the controlled search of the solution and eventually reaching the goal. The reduced forces in the field will also make moving against them easier.

Whether the individual employs these mechanisms of tension reduction depends on the level of tension in the PSR and the individual's ability to cope with this level of tension (i.e. whether or not the individual finds the level of tension to be 'distracting' or unproductive).

Due to the increasingly negative nature of the barrier, the force pushing away from it also increases. Consequently, the resultant of the forces acting in the field shifts in the direction away from the barrier and the person is "pushed out" from that sub-region. At this moment, the person might abandon the problem altogether and leave of the PSR to a neighbouring region. For example, one may decide to take a break from a problem and do something else. The PSR, however, still remains a region of high tension. If the new region and the PSR are segregated enough, the tension from the PSR will not affect the other region. The possibility of leaving the PSR depends on the strength of the boundaries of that region and the extent to which they are passable. For example, to physically leave the PSR region in a context of the experimental situation might not be possible without moving to the undesirable failure sub-region first. In this situation, the individual is more likely to remain in the PSR longer than they might have otherwise.

Alternatively, in response to increasing tension, the field of the PSR might be restructured such that locomotion to a different sub-region within the PSR that is not connected with the barrier becomes possible, or such that a new set of possible sub-regions appear. The locomotion to a different sub-region will lead to a structural change in the PSR. This change could merely be in the form of a mere differentiation of the unstructured space, a construction of a new sub-region, or it might involve a restructuring of the existing sub-regions, including the sub-region of the initial state and/or the goal state.

For example, one might notice a pattern in the values of the equation – 4, 3, 1 – and realise that these values could be engaged in three valid arithmetic expressions without changing the values themselves: (4 = 3 +1), (4 – 3 = 1), (4 – 1 = 3). The relationships among the values that are seen now are different from the ones constructed previously (e.g. 4 is the difference between the other two values). This, in turn, leads to a restructuring of the initial problem representation. For the solution of an insight problem, the restructuring of the psychological field is necessary. If the goal is not attainable from the new sub-region, the person encounters a new barrier and the process described above repeats, leading to further increases in the tension within the system.

When locomotion within the PSR results in the restructuring of the psychological field such that a path exists between the initial state and the goal state, the problem is finally solved. Restructuring the psychological field involves the change and reorganization of the previously constructed sub-regions of the initial state and the goal state of the problem within the PSR such that they become interdependent parts of a single whole. This reorganization involves changing the structure of these sub-regions in the grouping and relationships of their elements; that is, a restructuring of the problem representation. For example, the solution move in the discussed example, IV = III – I, involves in removing one stick from the equality sign and placing it next to the minus sign, turning it into "=." Thus, the solution is IV – III = I. The changes in the two elements of the initial equation, "=" and "–," also lead to relational changes among other elements in the equation. In this case, the element "IV" is no longer the outcome of the operation, but the element "I" is. As the person moves to the goal sub-region, the tension accumulated in their life space during the problem solving process is rapidly released.

As a result, the restructuring of the region of activity is accompanied by a rapid release of tension at the moment when the problem is solved. The rapid release of tension in insight problems is due to the fact that the solution to these problems is not attainable in a gradual manner. The fact that the goal is attained suddenly, all at once, and not gradually in insight problem solving was demonstrated by the studies of Metcalfe and Wiebe (1987) and Davidson (1995), and it is a commonly accepted characteristic of insight problems (e.g. Ohlsson, 1992).

The locomotion towards the goal in different situations allows releasing tension in the system to the extent that this locomotion is perceived as bringing the person closer to the goal. In situations where gradual progress to the goal is possible, the release of tension is also gradual. In these situations, the path to the goal becomes structured as a series of sub-goals, and reaching these sub-goals allows for a certain amount of tension to be released. Thus, the total tension in the system gradually decreases before the goal region is reached. For example, in the tower of Hanoi problem some tension is released with every move of a peg that takes an individual closer to the goal. As a result of a sequence of these moves, the goal region is approached step by step and, at the same time, the tension is released gradually via the locomotion through the series of sub-goals.

Before reaching the goal, the locomotion within the PSR itself partially releases some of the tension in the system. However, only the release of a minor portion of overall tension is possible through the locomotion within the PSR in insight problems. In an insight problem, locomotion through the PSR does not lead to a gradual arrival at the goal. The attainment of the goal is a natural way of releasing the tension associated with the need to solve the problem and in insight problems it is achieved in one single psychological step. Unless the locomotion leads to the goal sub-region in an insight problem, it will result in encountering yet another barrier, and result in a subsequent increase in tension. However, when the goal sub-region is finally reached, the tension accumulated in the PSR is released, resulting in a great amount of tension release in a single psychological step.

2.2.5. Characteristics of tension in the insight problem solving

The presence of barriers on the path toward the goal and the problem solver reaching actual state of conflict (impasse) are the characteristics of insight problems that contribute to an overall increase of tension in the whole PSR. In insight problems, very little tension is released through locomotion within the PSR unless the goal sub-region is reached, because gradual progress toward the goal is not possible. Thus, locomotion, if not successful, leads to a barrier and a subsequent increase in tension.

As a result, an insight problem region is characterised by a tense system that can only be relieved by entering the goal sub-region. Non-insight problems are characterized by a tense system that is released gradually through locomotion within the PSR in the direction og the goal and the attainment of sub-goals. At the moment the goal is reached in insight problems, all the tension is released rapidly at once. It is worth noting that it is not the amount of tension within the PSR that distinguishes it from non-insight problems, but rather the mechanism of its release.

The above analysis shows that the tension in the PSR in insight problem solving generally sustains itself and even increases until the solution is found. Encountering barriers and reaching a state of conflict in the field (impasse) contribute to the increase of tension in the system. Since neither the partial attainment of the goal, nor gradual progress toward it, is possible in insight problem solving, there are no mechanisms to substantially reduce the tension in the region other than solving the problem. The increase of tension might be controlled by adjusting one's cognitions with respect to the problem and its environment. However, this will most likely stabilise the level of tension rather than reduce it.

As a result, the tension accumulated in the process is alleviated in a single release when the problem is solved. A greater amount of tension at the moment of solution signifies a higher level difficulty experienced when looking for the solution. Thus, the following hypothesis is formulated.

Hypothesis 2: When the solution to the same problem is found with different levels of tension, the intensity of the insight experience will be greater in the case when the greater amount of tension is released with the solution of the problem.

To summarize, when an insight problem is being solved, depending on the dynamics of the problem solving process (e.g. reaching barriers and a state of conflict) a certain amount of tension is generated in the life space. When the problem is finally solved, the representation of the problem undergoes a certain amount of change, which is accompanied by a rapid release of the tension generated during the problem solving process. Both the degree of restructuring and the amount of tension released upon finding the solution contribute to the intensity of the insight experience.

2.3. Theoretical propositions

The insight problem solving process is characterized by the following distinctive properties:

- All change in representation necessary to solve the problem happen as a result of one psychological step;
- The amount of restructuring achieved over one step is significantly greater than in non-insight problems;
- Before the problem is solved, the amount of tension is either increasing or is sustained with no gradual tension reduction mechanism being available;
- At the moment of solution the tension from the problem solving region is released rapidly;
- The proportion of tension in the PSR released upon finding the solution move is significantly greater than in non-insight problems.

It was shown in the first theoretical section that the restructuring of the problem representation per step can vary among the situations. The dynamics of the solution process, such as the strength of the forces in the field and the various barriers encountered, might lead to the different levels of tension in the problem solving region. As a result, both the degree of restructuring and the level of tension can vary greatly in their value in different problem solving situations, thus leading to differences in the experience of problem's difficulty and the experience of insight.

It is proposed that the insight experience refers to the abrupt restructuring of the psychological field accompanied by a rapid release of tension. It was hypothesized that the insight experience, namely its intensity, varies depending on the two factors that determining it:

- The degree of restructuring involved in a problem, and
- The amount of tension released within the system when the problem is solved.

The intensity of the insight experience can be represented as a function of these two factors.

$$iI = F\ (R, T)$$

The intensity of the insight experience (iI) is a function (F) of the degree of restructuring (R) of the field and the amount of tension in the field (T) released at the moment of solution.

The degree of restructuring and the amount of tension are not entirely independent factors. It is reasonable to suppose that problem requiring a greater restructuring might lead to more frequent encounters with a barrier that will also result in a greater amount of tension being generated within the PSR. However, the degree of restructuring by itself does not entirely determine the amount of tension that will be generated during the problem solving process and later released at the moment of solution. There are other factors that affect the amount of tension generated in the PSR, such as the presence of an induced force and the magnitude of the valences associated with finding the solution and not finding it.

The exact nature of the functional relationship between the intensity of the insight experience, the level of restructuring, and the amount of tension has not been determined. However, both restructuring and tension are positively related to the intensity of the insight experience. This means that an increase/decrease in the amount of restructuring will correspond to an increase/decrease in the intensity of the insight experience given the same level of tension in the system; and an increase/decrease in the amount of tension released at the moment of solution will correspond to an increase/decrease in the intensity of the insight experience given the same degree of restructuring required for the solution. These interdependencies were captured in Hypotheses 1 and 2.

3. Method

3.1. Preliminary investigation

A preliminary investigation was conducted in an attempt to get a better understanding of people's experience during insight problem solving, and to get firsthand verbal descriptions of people's experience when they work on a problem and when they solve it. More specifically, evidence was sought for the descriptions of experience of tension before the solution and a description of the insight experience. Participants in this study were given a matchstick arithmetic problem to solve (Knoblich et al., 1999) and after solving it they were asked to describe their experiences before the solution and at the very moment of solution itself.

3.1.1. Method

Participants. Participants in this investigation were 20 undergraduate students who were offered a partial course credit for their involvement in the study.

Materials. Matchstick equation puzzles were used as stimuli (Knoblich et al., 1999). Several problems were used randomly, including IX = III + I (solution IV = III + I), IV = III – I (solution IV – III = I), and II – I = VI (solution II – I = I / I). They were represented with brown coffee sticks on the table where the participants were seated.

Procedure. The study was conducted in a laboratory setting with one participant at a time. Each session lasted from about half-an-hour to an hour. After arriving at the laboratory and signing the consent forms, the participants received informal verbal instructions about the purpose of the study and the procedure. Participants were told that with this study we tried to get a better understanding of people's experience when they work on and solve problems. The participants were told that they will be asked to solve a puzzle, and after they had solved it, they will be asked to describe their experiences before the solution and at the solution moment specifically. Then, participants received training in matchstick equations and Roman numeral representation. A list containing Roman numerals from 1 to 12 and five mathematical operations was available to the participants at all times during this study. This list is presented in Appendix B. After the training, each participant was given a matchstick equation to solve. There was no time constraint for participants to work on the problems. They were instructed that there was no consequence if they did not solve the problem, they could stop working on it whenever they wished. After they solved the problem, the participants were encouraged to freely describe their experience before they found the solution and at the solution moment itself in their own words. Each session was videotaped.

3.1.2. Results

All participants in this investigation solved the assigned matchstick equation. Participants' descriptions pertaining to their experience before the solution and at the moment of solution were transcribed. These descriptions were analysed with respect to comments

- describing experience similar to the concept of tension;
- pertaining to changes in tension during the problem solving process;
- with respect to release of tension after the solution;
- describing the solution moment.

The transcribed descriptions were coded using QSR N6 qualitative data analysis software (QSR International Pty Ltd, 2002).

3.1.2.1. Experience before the solution

All 20 participants mentioned experiencing a feeling that could be classified as either descriptions (or synonyms) of tension itself (e.g. pressure, stress) or the consequences of its increasing level (e.g. nervousness, frustration). However, none of the participants used the word "tension" per se. The descriptor words used by the participants and their frequencies are reported in Table 2.

Table 2: Descriptor words of the experience before the solution was found

Descriptor word	# of participants mentionned*	% of participants*
Stress	6	30%
Frustration	5	25%
Nervousness	5	25%
Pressure	5	25%
Uncertainty with respect to one's ability to solve the problem	3	15%
Panic	2	10%

*Some participants used more than one descriptor word, therefore the column sums are greater than 20 or 100%

Fourteen out of 20 participants (70%) explicitly described an increase in intensity of their feeling before the solution which was attributed to a passage of time and failing to find the right move.

Six participants (30%) mentioned some kind of tension reduction actions during the problem solving. However, these tension reduction activities were reported to take place in response to increasing tension to the point of being destructive. The following comment from one participant's description illustrates the idea:

> I tried a few things and they didn't work, so I got stressed because I couldn't really get it, and after a while it got to a peak, and I had to get relaxed and think what I am going to do and calm myself down…

Interesting enough, the participants who reported that they had to "calm" themselves down and get "relaxed" could not really explain how exactly they did it. One participant mentioned that he calmed himself down by saying to himself that "I can do this, it should be easy."

3.1.2.2. Experience at the moment of solution

In their description of the moment of solution, 16 participants (80%) commented on the suddenness property of finding the solution to their matchstick equation. Suddenness of the solution was illustrated in three ways:

- through directly using the word "sudden"
- commenting on the surprise factor of the solution
- using an exclamation indicating a surprise such as "WOW!" "BOOM!" or "Oh, yeah!"

In their description of the moment of solution, none of the participants used the word "insight" directly; only one person used the word "Eureka," and only one individual described it as an "AHA!" experience. However, the moment of solution descriptions of the participants did indicate an experience of a sudden realisation of the solution and the experience of insight. The descriptor words used by the participants that pertain to the indication of the experience of insight and their frequencies are reported in Table 3. The excerpts from participants' descriptions of their moments of solution in this study are reported in Appendix C.

Table 3: Solution moment descriptor words

Descriptor	Frequency	%	Descriptor	Frequency	%
Jolt/Peak of excitement	4	20%	AHA!	1	5%
Surprise/ WOW!/Oh yeah!	4	20%	Eureka	1	5%
Elation	3	15%	I got it!	1	5%
Epiphany	2	10%	Intense happiness	1	5%
Euphoria	2	10%	It clicked	1	5%
Genious/triumphant moment	2	10%	Spark	1	5%
Rush	2	10%			

The accounts of the moment of solution experience reported by the participants indicated that the solution of an insight problem had an emotional effect besides cognitive realisation of the solution itself.

The most commonly used descriptor of the moment of solution was the word “relief.” Fifteen participants (75%) commented that finding the solution to the problem was associated with a feeling of “relief.” One participant described her experience when she solved the problem as “I could feel the pressure drop, I could feel my shoulders going from being tense and up to down and relaxed. Relief was a big feeling…”

Other feelings, such as excitement (50%), happiness (40%), success (25%) and pride (15%) were reported as well indicating an overall emotionally positive sensation resulting from solving the problem.

Two participants commented on the relationship between the intensity of their solution moment (“the rush” and “the spark”) and the amount of frustration and time spent solving the problem. Both participants indicated that the longer one spends solving the problem, the more intense their solution experience would be. This relationship is captured in Hypothesis 2 and tested in Experiments 3 and 4.

3.1.3. Discussion

Participants in this study were invited to freely describe their experience during the process of finding the solution to an insight problem and at the very moment of solution itself. One-move matchstick equations (Knoblich et al., 1999) were used as insight problems in this study. Specifically, an evidence of the experience of tension during the problem solving and a description of the insight moment itself were sought.

The obtained descriptions showed that tension is a phenomenologically valid experience during problem solving. It was mentioned by all of the participants in one form or another. More than half of the participants (70%) volunteered a description of an increasing trend in this feeling before the problem was solved, and 75% of the participants associated the solution to the problem with a “relief” and “pressure drop.” The commonly noted feeling of “relief” at the moment of solution directly points to the release of tension.

These descriptions confirm the theoretically suggested dynamics of tension during the insight problem solving. It was argued that tension is either increasing or is sustained during the solution process, and that it is released after the solution is found. It is worth noting, that all these descriptions were provided in response to the open-ended question: “Could you please describe your experience before you found the solution to the problem?” The obtained descriptions were volunteered by the participants, because the purpose of the investigation was to get people’s firsthand description of their experience without suggesting the answers to them. More pointed questions during the interview might have resulted in a higher frequency of the accounts of increasing tension trend and a feeling of relief after the solution.

Moreover, quite vivid descriptions of the solution moment experiences were obtained from the participants. It is interesting to note that the word “insight” itself was not used at all in the participants’ descriptions. It seems that this word is more commonly used by people to mean knowledge and understanding[2] similar to Smith’s (1995) definition, as opposed to the “insight experience” referring to “the sudden emergence of an idea into conscious awareness, the “Aha! experience” (Smith, 1995, p. 232). In their descriptions, the participants indicated a strong, positive emotional connotation of the solution moment. In fact, the participants’ accounts for the solution process suggested that an emotional switch happened when an insight problem was solved: it was a switch from a feeling of pressure and stress to excitement and happiness accompanied by a relief.

---

[2] In the Oxford English Dictionary (1989) the word “insight” is defined as

a. Internal sight, mental vision or perception, discernment; in early use sometimes, Understanding, intelligence, wisdom.
b. Knowledge of or skill in a particular subject or department

3.2. *Experiment 1*

The purpose of Experiment 1 was two-fold. First, it was designed to test Hypothesis 1 by assessing the relationship between the degree of restructuring in a problem and the intensity of the insight experience. Hypothesis 1 states that, when two problems are solved under the same conditions, the intensity of the insight experience will be greater for the problem that requires the greater degree of restructuring. In addition, the data collected in Experiment 1 was used to obtain initial empirical support for the developed analytical measure of restructuring. It was predicted that in each pair, the problem with a higher change score computed in advance will be judged to be the more difficult of the two. The degree of restructuring was varied across problems in a within-participant design.

3.2.1. Method

Participants. 155 undergraduate students from the University of Waterloo participated in the study in exchange for a partial course credit.

Materials. Ten matchstick equations with varying change scores were used in this study as stimuli. Table 4 shows the equations with their solutions and corresponding change scores.

Table 4: Ten matchstick equations used in Experiment 1

Type	#	Problem	Solution	Change Score
Value	1	VII – II = III	VI - III = III	2
	2	IV – III = III	VI – III = III	3
Tight Chunk	3	V + V = V	V + V = X	5
	4	IX = III + I	IV = III + I	7
Operation	5	III - II = IV	II + II = IV	9
	6	VI + II = III	VI – III = III	10
	7	IV = III – I	IV – III = I	11
Tautology	8	III + III = III	III = III = III	12
	9	III – II = II	II = II = II	12
Division	10	II – I = VI	II – I = I / I	19

Ten problems could be paired among each other yielding a set of 55 possible different pairs. A set of 24 pairs of equations was selected from a possible set of 55 pairings of the ten equations to be used as stimuli for this study. Not all possible pairings of the ten problems were used to keep the study manageable. The primary purpose of this study was testing Hypothesis 1, while the secondary goal was to provide an initial, preliminary test of the developed change measure. The developed change measure does not constitute a major part of this work, and therefore, given the time and resource constraints, only preliminary testing was attained. The selected 24 stimuli pairs represented a variety of change score comparisons, including within-type comparisons (based on problem types proposed by Knoblich and colleagues (1999)), comparisons of tight-chunk problems to problems of other types, comparisons of division problems to other types, and between-type comparisons. Table 5 shows the 24 selected pairs of problems with their solutions and change scores organized into the above categories.

Procedure. Participants received an interactive Portable Document Format (PDF) form by e-mail. They were instructed to fill out the form in one session and return their responses by e-mail. 182 students volunteered to participate in the study; 182 forms were sent out and 155 responses were returned. Each interactive form consisted of 17 pages containing a consent form, set of instructions and 12 experimental pages with six pairs of matchstick problems. There were two pages for each pair of matchstick problems. On the first page, the two problems were presented side by side with their solutions below. The participants were asked to evaluate the solutions to the problems and indicate on 10-point scales the difficulty and insightfulness of each problem in each pair as it might have been judged by someone who actually solved the problems. On the second page for each pair, participants were asked to provide comments to help understand their responses. Providing comments was optional. Each participant received a randomly selected and arranged set of six pairs of matchstick problems. The two problems within each pair were presented in a random order as well. After completing their form, participants were instructed to return their responses by e-mail. The completed or partially completed form could not have been saved, which enforced completion of the experimental task in a single session as opposed to several sessions. The responses were returned in the Extensible Markup Language (XML) format. Appendix D presents a sample of the document sent to the participants.

Table 5: The 24 pairs of matchstick problems used in Experiment 1

Pair	P1 (Problem 1)	P1 Change Score	P2 (Problem2)	P2 Change Score	Score Difference (P2-P1)
			within group differences (6)		
1	VII – II = III	2	IV – III = III	3	1
2	V = V + V	5	IX = III + I	7	2
3	III - II = IV	9	VI + II = III	10	1
4	III - II = IV	9	IV = III – I	11	2
5	VI + II = III	10	IV = III – I	11	1
6	III + III = III	12	III – II = II	12	0
			positionning X – V transformation (6)		
7	IX = III + I	7	VII – II = III	2	-5
8	IX = III + I	7	IV – III = III	3	-4
9	IX = III + I	7	III - II = IV	9	2
10	IX = III + I	7	VI + II = III	10	3
11	IX = III + I	7	IV = III – I	11	4
12	IX = III + I	7	III + III = III	12	5
			positionning / problems (5)		
13	IV – III = III	3	II – I = VI	19	16
14	IX = III + I	7	II – I = VI	19	12
15	VI + II = III	10	II – I = VI	19	9
16	IV = III – I	11	II – I = VI	19	8
17	III + III = III	12	II – I = VI	19	7
			Between group comparisons (7)		
18	VII – II = III	2	VI + II = III	10	8
19	IV – III = III	3	III + III = III	12	9
20	VI + II = III	10	III + III = III	12	2
21	IV = III – I	11	III + III = III	12	1
22	IV – III = III	3	V + V = V	5	2
23	III – II = II	12	II – I = VI	19	7
24	IV = III – I	11	III – II = II	12	1

Measures. The degree of restructuring was measured in two different ways. It was measured as a change score computed in advance using the procedure outlined in Section 2.1.3 above. In addition, the degree of restructuring was measured as a relative judgment of difficulty of the two problems in a comparison pair, which was obtained on a 10-point scale for each problem in a pair. In the experimental task, problems were presented along with their solutions; the participants did not have to solve the problems by themselves, but were instead asked to appreciate and evaluate the provided solutions. The participants had a fair degree of flexibility with respect to where and when to complete the task. Presumably, a more convenient time and place was chosen for the completion of the task. As a result, the experimental task was completed at a time and location chosen by the participant, and required only the evaluation of solutions to problems rather than solving them. In such an experimental situation, the levels of tension associated with actually having to solve the problems, or being observed while completing the task, were minimized. Under these conditions, the perception of relative difficulty was mainly based on differences in the difficulty levels of transformations from the initial states of the problems to their goal states, with little if any effect of tension. Therefore, it is reasonable to assume that the perception of the relative difficulty of the two problems in each comparison pair primarily reflected the difference in the perceived difficulty of the transformation from the initial state to the goal state of the two problems. A more difficult transformation requires more change from the standpoint of the judging individual, or in other words, a greater degree of restructuring. The subjective judgment of difficulty of the problems in pairs was used as an independent measure of the degree of restructuring since the effect of other factors potentially contributing to the perception of difficulty of a problem was minimized.

The intensity of the insight experience was measured as a relative judgment of “insightfulness” of the two problems in a pair, which was obtained on a 10-point scale for each problem in the pair. Insight was defined as an “Aha! I see now!” experience; as an experience of a "light bulb flashing" above one’s head. The definition and instructions for providing this judgment were presented on page 4 of the experimental form (see Appendix D).

3.2.2. Results

The purpose of Experiment 1 was two-fold: To test Hypothesis 1, and to collect an initial set of data for the preliminary verification of (and possible adjustment of) the developed measure of restructuring. The results are reported separately for each of the two objectives.

3.2.2.1. Testing Hypothesis 1

The following two experimental predictions derived from Hypothesis 1 were tested:

1. When comparing two problems, a problem that was judged to be the more difficult of the two is more likely to also be judged as the more insightful problem of the pair.
2. There will be a positive correlation between the difference in the judgment of the difficulty of the two problems and the difference in the insight judgment of the two problems in a pair.

Insight judgments for the two problems in the first comparison pair of each participant were analyzed to establish whether a judgment of higher difficulty corresponded to a judgment of higher insight. This resulted in a set of 155 pair-wise comparisons randomly drawn from the set of 24 pairs of matchstick equations presented in Table 5. For each pair, the prediction was that a problem that was judged to be more difficult by a participant will also be judged to be more insightful. The pairs where both problems received identical scores of difficulty were excluded from this analysis ($N_{D1=D2}$ = 16). Table 6 presents the medians and quartiles of the insight judgment for the problems that were perceived as more difficult and less difficult in each pair.

Table 6 shows that problems that were perceived to be less difficult were also perceived to be less insightful. Since the judgments of insight were obtained on an ordinal scale, a Wilcoxon signed ranks test was computed to assess the statistical significance of the observed difference in the judgment of insight. Problems that were judged as being the more difficult in each pair were also judged to be significantly more insightful than problems that were judged as being the less difficult of the two ($Z = -4.30$, $p < 0.001$). The difference in insight judgment predicted by experimental prediction 1 was statistically significant, thus providing support for Hypothesis 1.

Table 6: Judgment of insight for the two problems in the first comparison pair (*N*=139)

	Problem perceived as less difficult in a pair	Problem perceived as more difficult in a pair
Lower *Q*	3	5
Mdn	4	7
Upper *Q*	6	8

As a second step in testing hypothesis 1 through experimental prediction 2, the differences in the judgments of difficulty and insight of the two problems in each comparison pair were computed. Spearman's rho correlation coefficient between the difference in the judgment of difficulty and the difference in the judgment in insight was 0.614 ($N = 155$, $p < 0.01$). The presence of a positive correlation between the difference in perceived difficulty and the difference in perceived insight indicates that a larger difference in difficulty is also associated with a larger difference in insight judgment. This supports experimental prediction 2 and Hypothesis 1.

3.2.2.2. Verifying the measure of restructuring

Based on participants' comments and a careful examination of the problems, three pairs were removed from the set for further analysis with respect to verifying the change score: pairs 2, 18, and 22. In the case of pair 18 (see Table 5), the two problems had different initial states and required quite different transformations to be performed to solve the problems, but both problems had identical solution states (VI – III = III). This peculiarity caused confusion among the participants, who often commented on this issue. Less than 54% of the participants ($N_{P18} = 39$) for this pair provided comments, and 33% of those who provided comments (or 18% of all the participants in this pair) clearly commented on the similarity of the two solution states. These participants based their judgment on the end state similarity of the two problems as opposed to examining the structural transformations required to achieve that state in each of the problems. For the remaining participants who provided comments, it was not always possible to determine whether or not their judgment was influenced by this coincidence since comments were often very brief and did not always fully explain the reasoning behind the judgment. The reasoning of the remaining participants who chose not to leave a comment at all was completely unclear. In this case, the decision was made to exclude pair 18 from the validation of change score analysis as the change score is based on the transformation of the initial state to the goal state, and not on the similarity of the two end states.

The pairs 2 and 22 both involved the problem V + V = V with the provided solution V + V = X. This problem could also be solved as a "tautology" with the solution V = V = V. These two solutions to the same problem result in two different change scores. The former solution results in a score of 5, while the tautology solution to this problem produces a score of 12. The tautology solution to this equation proved to be so powerful that 25% of the participants in pairs 2 and 22 who provided comments clearly remarked on this property of the problem. The comments of the participants were often very short and vague, and it was sometimes impossible to judge the exact properties of the problem this particular individual had paid attention. The effect of the alternative solution on the judgment of difficulty of this problem in relation to the other problem in the pair was also difficult to determine. Some participants commented that the fact that there are two different solutions to this problem makes it seem easier, while other participants stated that this makes the problem more difficult. As a result, it was not clear which solution to the problem V + V = V each given participant evaluated (even if they provided comments). Even when the participants were aware of the alternative solution, it was also not clear how it affected their judgment.

As a result, due to the ambiguity in the stimuli, pairs 2, 18, and 22 were excluded from the validation of the measure of restructuring.

Experimental predictions for validating the change score were the following:

1. In each pair, the problem that has a higher change score will be judged to be the more difficult of the two problems;
2. There will be a positive correlation between the difference in the change scores of the two problems in a pair and the average difference in the subjective judgments of difficulty of these problems;

Table 7 presents descriptive statistics for the judgment of difficulty for each problem (one with a higher change score and one with a lower change score) in each pair. While there was a noticeable difference between the difficulty judgments of the two problems in some pairs, in other pairs it was not so pronounced.

Single Wilcoxon Signed Ranks tests were computed to assess the significance of the difference in difficulty judgment in each pair between problems with higher and lower change scores. A procedure for controlling the False Discovery Rate (FDR) (Benjamini & Hochberg, 1995; Benjamini & Yekutieli, 2001) was used in determining the significance of the individual test results in a multiple inference situation. Table 8 presents results of the tests along with the assessment of the significance of the results while controlling the FDR at 0.05 and 0.06 levels. Based on the test results reported in Table 8, there were:

- 13 pairs with significant differences in the judgments of difficulty at 0.05 FDR level in the predicted direction;
- two pairs with marginally significant differences (controlling FDR at 0.06 level) in the judgments of difficulty in the predicted direction (pairs 5 and 8);
- three pairs with significant differences observed in the opposite direction from the predicted direction (pairs 9, 10, and 11);
- no significant difference in the judgment of difficulty was predicted and observed in pair 6;
- no significant differences in the judgment of difficulty was observed in pairs 12 and 21, although the presence of differences was predicted.

Table 7: Descriptive statistics for the judgment of difficulty for the problem with lower and higher change score in each pair analyzed

Pair		N	Percentiles		
			25th	50th (*Median*)	75th
1	Problem with lower change score	34	1.75	3	6
	Problem with higher change score		3	4.5	7
3	Problem with lower change score	37	2	4	5
	Problem with higher change score		3	5	7
4	Problem with lower change score	45	2	4	6
	Problem with higher change score		3.5	5	7
5	Problem with lower change score	41	3.5	5	7
	Problem with higher change score		3.5	6	8
6	Problem with lower change score	37	2	4	7
	Problem with higher change score		2	5	7.5
7	Problem with lower change score	40	2	4	5
	Problem with higher change score		4	6.5	8
8	Problem with lower change score	40	3	5	6
	Problem with higher change score		4	6	7
9	Problem with lower change score	36	4	8	9
	Problem with higher change score		3	5	7
10	Problem with lower change score	36	5.25	7	8
	Problem with higher change score		3.25	5	7
11	Problem with lower change score	37	3.5	6	8
	Problem with higher change score		2.5	4	7
12	Problem with lower change score	38	3	6	7
	Problem with higher change score		3.75	5	7
13	Problem with lower change score	37	2	4	6.5
	Problem with higher change score		5.5	8	9
14	Problem with lower change score	41	3	5	7
	Problem with higher change score		6	7	8
15	Problem with lower change score	36	2	4	6
	Problem with higher change score		5	6	9
16	Problem with lower change score	41	3	5	7
	Problem with higher change score		5.5	7	8
17	Problem with lower change score	39	2	5	7
	Problem with higher change score		4	7	9
19	Problem with lower change score	37	2.5	4	6
	Problem with higher change score		3	6	7
20	Problem with lower change score	37	3	5	6.5
	Problem with higher change score		4	7	8
21	Problem with lower change score	37	3	5	7
	Problem with higher change score		4	5	8
23	Problem with lower change score	44	4	6	8
	Problem with higher change score		6	7	8
24	Problem with lower change score	41	4	5	6
	Problem with higher change score		5	7	8

Table 8: Results of single Wilcoxon Signed Ranks tests computed for 21 pairs along with the significance assessment for each test result based of the FDR procedure at 0.05 and 0.06 levels

Pair #	Change Score P1	Change Score P2	Direction of the predicted difference	Direction of the observed difference	*N*	*Z*	*p* (1-tailed)	FDR at q* = 0.05	FDR at q* = 0.06
1	2	3	P1 < P2	P1 < P2	34	-3.36	< 0.001	Sig.	Sig.
3	9	10	P1 < P2	P1 < P2	37	-3.27	0.001	Sig.	Sig.
4	9	11	P1 < P2	P1 < P2	45	-2.16	0.015	Sig.	Sig.
5	10	11	P1 < P2	P1 < P2	41	-1.64	0.050	Not Sig.	Sig.
6	12	12	ND	ND	37	-0.86	0.390**	Not Sig.	Not Sig.
7	7	2	P1 > P2	P1 > P2	40	-4.74	< 0.001	Sig.	Sig.
8	7	3	P1 > P2	P1 > P2	40	-1.65	0.049	Not Sig.	Sig.
9	7	9	P1 < P2	P1 > P2	36	-2.82	0.005**	Sig.	Sig.
10	7	10	P1 < P2	P1 > P2	36	-3.28	0.001**	Sig.	Sig.
11	7	11	P1 < P2	P1 > P2	37	-2.04	0.037**	Sig.	Sig.
12	7	12	P1 < P2	ND	38	-0.81	0.208	Not Sig.	Not Sig.
13	3	19	P1 < P2	P1 < P2	37	-4.39	< 0.001	Sig.	Sig.
14	7	19	P1 < P2	P1 < P2	41	-3.83	< 0.001	Sig.	Sig.
15	10	19	P1 < P2	P1 < P2	36	-4.26	< 0.001	Sig.	Sig.
16	11	19	P1 < P2	P1 < P2	41	-3.46	< 0.001	Sig.	Sig.
17	12	19	P1 < P2	P1 < P2	39	-3.06	0.001	Sig.	Sig.
19	3	12	P1 < P2	P1 < P2	37	-1.82	0.034	Sig.	Sig.
20	10	12	P1 < P2	P1 < P2	37	-2.33	0.010	Sig.	Sig.
21	11	12	P1 < P2	ND	37	-0.76	0.225	Not Sig.	Not Sig.
23	12	19	P1 < P2	P1 < P2	44	-2.12	0.017	Sig.	Sig.
24	11	12	P1 < P2	P1 < P2	41	-2.46	0.007	Sig.	Sig.

* *q** is a False Discovery Rate significance level
** two-tailed *p* value

The overall prediction rate based on the change score computed in advance was 66.7%. Accepting a marginal significance level of FRD at 0.06 for pairs 5 and 8, the observed prediction rate increases to 76.2%.

There was a significant correlation between the difference in the change scores of the two problems in the pairs and the average differences in the judgment of difficulty (Spearman's $\rho = 0.50$, $N = 21$, $p < 0.01$) and a significant correlation between the difference in the change scores of the two problems in the pairs and the average differences in the judgment of insight in the pairs (Spearman's $\rho = 0.37$, $N = 21$, $p < 0.05$).

The presence of a positive correlation between the difference in the change scores and the average differences in the judgment of difficulty and insight indicates that an increase in the difference in the computed change scores of the two problems was also accompanied by an increased difference in the perception of difficulty of the two problems, as well as an increased difference in the insight judgment of the same two problems.

3.2.3. Discussion

3.2.3.1. Possible adjustment to the change score

In pairs 9, 10 and 11, the difference in the difficulty judgment was observed in the opposite direction from what was predicted. All three of these pairs involved a comparison of a “tight chunk” problems (Knoblich et al., 1999) - where an X symbol had to be changed into a V symbol - to problems involving various changes in operations which had different change scores (Table 5). In pair 12, no significant difference was observed between the judgments of each problem’s difficulty. In this pair a “tight chunk” problem was compared to a “tautology” problem, which involved a change in operation. In the study, the problem involving the X ↔ V transformation was consistently judged as being significantly more difficult than the various operation problems with change scores of 9, 10 and 11. There was no significant difference in difficulty judgments when compared to a problem with a change score of 12.

These results suggested that the computation of the change score for the X ↔ V transformation had to be adjusted to reflect the greater difficulty of such change as perceived by participants. To be consistent with the observed differences in difficulty judgments, the change score for problem number 4 (Table 4) could be adjusted from 7 units to around 12 units, implying that the weight of the X ↔ V transformation had to be at least doubled. If such a change was implemented, the differences in the change score would explain 85.7% of the observed differences in the judgment of difficulty levels. Furthermore, if one was willing to accept a marginal significance of FDR at a 0.06 level, 95.2% of the results could be explained by the difference in change scores.

A re-computation of the correlation coefficients reported above, taking into account the proposed change in the computation of the change score, led to an increase in the coefficients' values and their significance levels: The re-computed correlation coefficient between the difference in the change score and the average differences in the judgment of difficulty (Spearman's $\rho = 0.64$, $N = 21$, $p < 0.001$); and re-computed correlation coefficient between the difference in the change score and the average differences in the judgment of insight (Spearman's $\rho = 0.49$, $N = 21$, $p < 0.01$).

Although no difference in difficulty judgment was observed for pair 21 a highly significant difference was found for pair 24. In both of these pairs, the same "operation change" problem was compared to two different versions of a "tautology" problem. The difference in the change scores of the two problems in both pairs was 1 unit, which might not have been as easily detectable by participants. The discrepancy in the observed difference in difficulty judgments in these two pairs could be interpreted as an indication that the two tautology examples were not equal in their difficulty level. One of the reasons for this difference may be the presence of a high degree of symmetry in one problem (III + III = III) and the lack of it in the other problem (III – II = II). Another plausible contributor to this phenomenon could be the size of the search space of "legal" moves available in each problem, which is greater for the non-symmetrical problem and smaller for the symmetrical one. Both of these factors were not included in the computation of the change score in its current form. It is worthwhile exploring how the symmetry in a problem and the size of the search space could be incorporated into the computation of the change score in the future.

3.2.3.2. Evaluation of the results with respect to taxonomy of matchstick problems proposed by Knoblich et al. (1999)

Table 9 provides a detailed, pair-by-pair comparison of the predictions and their accuracy made using the two different approaches: the computation of the change score and problem taxonomy.

The set of 21 pairs was selected for this study such that it allowed for a variety of comparisons, including within-type comparisons (based on problem types proposed by Knoblich and colleagues (1999)), comparisons of tight-chunk problems to other types of problems, comparisons of division problems to other types of problems, and between-type comparisons.

There were five within-type problem comparisons. In four of them, the difference in the judgment of difficulty was predicted, and in the fifth comparison no difference in difficulty was expected. Pair 1 compared two problems that were both classified as type A by Knoblich et al. (1999), yet received different change scores. Pairs 3, 4 and 5 compared three different problems of type B that also received different change scores. It is worth noting, that out of four within-type problem pairs, all four were observed in the predicted direction by the change score differences; three of them were significant at 0.05 FDR level and one was significant at 0.06 FDR level. It was assumed that in all the within-type problem comparisons, the problem taxonomy of Knoblich et al. (1999) would not predict a difference in difficulty. Although Knoblich et al. (1999) did not directly discuss and test the within-type differences. However, it is reasonable to assume that problems that were put in the same class would have had the same degree of difficulty.

In the pairs that compared tight chunk problem with other types - pairs 7 and 8 - the prediction was the same for both the taxonomy and the change score.

For the remaining four pairs that compared the X ↔ V transformation to different operation-type problems and tautology-type problems, the taxonomy of matchstick equations was unable to predict the difference in the difficulty judgment. The participants consistently judged the X ↔ V transformation problems to be harder than the operation problems, which was contrary to the direction predicted by the change score. These findings imply that the computation of the change score underestimated the difficulty of the X ↔ V transformation.

Table 9: Comparison of the experimental predictions made based on the change score versus the taxonomy of matchstick problems (Knoblich et al., 1999)

Pair	P1 change score	P2 change score	Change Score Difference	Change Score prediction	Knoblich et al. (1999) prediction	Observed direction	Change score predicted?	Knoblich et al. (1999) predicted?	Corrected change score
				Within category comparisons					
1	2	3	1	P1<P2	P1=P2	P1<P2	1	0	1
3	9	10	1	P1<P2	P1=P2	P1<P2	1	0	1
4	9	11	2	P1<P2	P1=P2	P1<P2	1	0	1
5	10	11	1	P1<P2	P1=P2	P1<P2*	1*	0*	1*
6	12	12	0	P1=P2	P1=P2	P1=P2	1	1	1
				Positioning X ↔ V transformation					
7	7	2	-5	P1>P2	P1>P2	P1>P2	1	1	1
8	7	3	-4	P1>P2	P1>P2	P1>P2*	1*	1*	1*
9	7	9	2	P1<P2	No Pred.	P1>P2	0	N/A	1
10	7	10	3	P1<P2	No Pred.	P1>P2	0	N/A	1
11	7	11	4	P1<P2	No Pred.	P1>P2	0	N/A	1
12	7	12	5	P1<P2	No Pred.	P1=P2	0	N/A	1
				Positioning division problems					
13	3	19	16	P1<P2	No Pred.	P1<P2	1	N/A	1
14	7	19	12	P1<P2	No Pred.	P1<P2	1	N/A	1
15	10	19	9	P1<P2	No Pred.	P1<P2	1	N/A	1
16	11	19	8	P1<P2	No Pred.	P1<P2	1	N/A	1
17	12	19	7	P1<P2	No Pred.	P1<P2	1	N/A	1
				Between catogory comparisons					
19	3	12	9	P1<P2	P1<P2	P1<P2	1	1	1
20	10	12	2	P1<P2	P1<P2	P1<P2	1	1	1
21	11	12	1	P1<P2	P1<P2	P1=P2	0	0	0
23	12	19	7	P1<P2	No Pred.	P1<P2	1	N/A	1
24	11	12	1	P1<P2	P1<P2	P1<P2	1	1	1
						Correct, # of pairs	16	6	20
			Able to predict, # of pairs	21	11	from predicted, %	76.19%	54.55%	95.24%
			Able to predict, % of the total set	100%	52.38%	from total, %	76.19%	28.57%	95.24%

* Significant with FDR at 0.06 level

There were six pairs of problems in the set that investigated the comparison of a division problem to a variety of other problems from different types with different change scores. It was unclear where the division problem was to be placed in the taxonomy of Knoblich and colleagues, as division and multiplication operations were not used in their set of problems. In terms of constraints, a division problem involves a change in both a 'tight chunk' and, to some degree, an operator. In the division problem, the existing operator is not changed, but a new operator is created. The division problem does not involve the tautology constraint per se; the solution to the problem changes the initial structure of the equation by creating a new number and a new operator. The division problem could not be classified as tautology, since it did not create a "meaningless" expression which is done when a second equality is constructed. Consequently, the division problem could be classified only as either a 'tight chunk' or an operator type, but it failed to fit the description of either category exactly.

Even if the division problem were classified as either a tight chunk or an operator, it would not result in greater prediction accuracy for the taxonomy of the six division problems. For example, if the division was classified as tight chunk, it would result in one correct prediction (pair 13), one incorrect prediction (pair 14) and four undetermined predictions (pairs 15, 16, 17, and 23). If the division problem were classified as an operator, it would result in one correct prediction (pair 13), four incorrect predictions (pairs 15, 16, 17, and 23) and one undetermined (pair 14).

Since the position of the division problem in the taxonomy was not exactly clear, a prediction for comparing the division problem to other types could not be made from the point of view of the taxonomy of problems.

In the between-category group of comparisons (pairs 19, 20, 21 and 24), predictions based on both the change score and the taxonomy were made in the same direction and, thus, had the same accuracy (three out of four pairs predicted correctly).

Overall, out of the set of 21 pairs used for the present study, the taxonomy of matchstick problems (Knoblich et al. 1999) could only be used to predict the differences in the difficulty judgment in 11 pairs (52.4% of the set), and predicted them correctly in only 6 pairs (28.6% of the set). When using the change score, however, the correct prediction was made in14 pairs, or 66.7% of the set (16 pairs, or 76.2% of the set with FDR at 0.06).

The computation of the change scores provided the advantage of having a single-dimensional, continuous scale measure for the degree of restructuring in matchstick equations. It allowed for the comparison of problems from the same type as well as problems from different types, and predicted the differences in situations where the taxonomy of Knoblich et al. (1999) was inapplicable or inaccurate.

Clearly, the results of Experiment 1 are only the first step in evaluating the proposed measure of restructuring. A more thorough analysis and empirical validation of this measure needs to be done to make any conclusive remarks. However, these results are encouraging and provide initial support for the change score measure.

3.2.3.3. Evaluation

Experiment 1 achieved two purposes: It provided both empirical support for Hypothesis 1 and also allowed for a preliminary evaluation of the developed measure of restructuring in the domain of matchstick equations.

It was crucial in this study (particularly for testing Hypothesis 1) to remove any and all possible sources of tension due to the experimental situation, so that the change in problem representation was the main influence in a problem's perceived difficulty level. Since psychological tension inevitably arises in a problem solving situation, the participants were not asked to solve the problems, and instead the problems were presented in pairs along with their solutions. The task of participants was to evaluate the solutions relative to each other and provide a relative rating of the two problems.

A judgment of the relative difficulty of two problems without having to actually solve them might not be exactly the same as a judgment from someone who in fact has solved the two problems. On the other hand, when an individual is engaged in the process of finding the solution, other factors aside from the restructuring affect the "experienced" difficulty of the problem. This makes it extremely difficult to isolate the impact of the degree of restructuring from other contributors to a participant's overall experience of a problem solving situation.

The participants completed their task outside of the experimental laboratory, and it follows that here was no control over the process of completing the task. The data, no doubt, has a certain degree of noise in it.

Given all these considerations, the set up of the experimental task, with its shortcomings and drawbacks, nevertheless served the purpose of creating a situation with minimum tension, which allowed examining the effect of the degree of restructuring in the domain of insight problems.

Since there might be an inconsistency between the judgment of insight based on a demonstration of a solution to a problem versus an experience of in fact finding that solution, another study was conducted where participants actually had to solve two problems and provided their judgment after they found the solutions themselves.

In the next experiment, Experiment 2, participants had to solve two problems of different degrees of difficulty and were asked for a relative judgment of the intensity of their insight on the two problems.

3.3. Experiment 2

The purpose of Experiment 2 was to investigate the effect of differences in the degree of restructuring on the insight experience in situations where participants had to solve the problems themselves. Under these conditions, the levels of tension in individual's life space could not be controlled since psychological tension unavoidably arises in a problem solving situation. The environment was kept constant for the two problems, and any differences in tension levels would naturally result from both the difference in difficulty levels of the two problems and the specifics of the structure of the life space of the individuals. Each participant solved a relatively easy problem followed by a relatively hard problem, and indicated the relative intensity of their insight experience on the two problems. This experiment utilized a within-participant analysis.

3.3.1. Method

Participants. 42 undergraduate students from the University of Waterloo participated in this study in exchange for a partial course credit. Only 28 participants solved both problems during the experimental session.

Materials. The stimuli for this study were the two matchstick problems with different degrees of restructuring, problem # 4 and problem # 10 from Table 4.

- Problem # 4: IX = III + I, solution IV = III + I
- Problem # 10: II – I = VI, solution II – I = I / I

These two problems were compared to each other in pair 14 of Experiment 1, and the results indicated that problem # 10 was judged to be significantly more difficult than problem # 4. These two problems were chosen as stimuli because there was a significant difference in their difficulty, yet both of them were not trivial.

The problems were set up in front of the participants on a table using brown coffee sticks for representation inside two taped together legal-size yellow folders. The sticks were placed on one side of the folder while the other side of the folder was used to cover the problem. When the folder was closed, the participant could not see the problem. The folder was located on the table in front of the participant. To start working on the problem, the participant opened the folder to reveal the puzzle. This allowed marking the beginning of the problem solving process.

Procedure. The study was conducted in a laboratory setting with one participant at a time. Each experimental session lasted one hour and was video taped using a recorder. After arriving at the laboratory and signing the consent forms, the participants received a set of instructions about the study in general, and matchstick equations in particular. The transcription of the instructions is provided in Appendix E. Care was taken to ensure that participants were comfortable with the laboratory setting. Participants were told that the study investigated what people experience when they solve problems and that they will be asked to solve two problems and answer several questions about their experience after they have solved the problem. Insight was defined as an “Aha! I see now!” experience, as an experience of a "light bulb flashing" above one’s head (see Appendix E for more details). Participants also received training in matchstick equation problems and the matchstick representation of the symbols. Each acceptable symbol (Roman numerals from 1 to 12 and five operations) was constructed with the coffee sticks on the table in front of the participant and its meaning was stated. A list containing Roman numerals from 1 to 12 and five mathematical operations was available to the participants at all times during this study. This list is presented in Appendix B. Participants had an opportunity to ask questions and clarify any uncertainties. After all the questions were answered, and the participant declared that they were ready to begin working on their first problem, problem # 4 was constructed inside the folder on the table. The participants could start working on the problem any time they were ready to begin by simply opening the folder and revealing the problem. Participants could work on the problem for as long as they wished, and no time constraint was imposed.

After a participant declared that they had solved the problem, the solution was verified and the participant was asked a series of questions. First, each participant received instructions and was asked to draw their tension curve for their first problem. To avoid making an assumption regarding participants' experience of insight, the participants were asked to report whether they had an experience of insight when they solved the problem (e.g. Bowden et al., 2005). If the answer was affirmative, the participant was asked to go back in their mind to the moment when they realized the solution and "get a good feel for that moment, especially its intensity."

The second problem, problem # 10, was set up inside the folder with the same instructions as the first one. Before a participant started working on their second problem, they were asked to remember their solution moment on the first problem one more time. The participants could start working on their second problem any time by simply opening the folder and revealing the problem. There was no time constraint for the second problem, but if a participant could not solve the problem within 20 minutes, they were stopped and were offered to have the solution shown to them.

After a participant solved the second problem, their solution was verified, and they were asked a series of questions. First, they were asked to draw their tension curve relative to that of the first problem. The participants drew both of their tension curves on the same sheet using different coloured markers. The participants were instructed to show any differences in their feeling of tension on the second problem compared to the first in their drawing using the distance to illustrate the proportionality of the difference. Participants were then asked to show the difference in the intensity of their insight moment between the problems on an unmarked scale. They were asked to position two points on the scale, representing the two solution moments, and position them such that the relative distance between the points and the beginning of the scale represented the difference in the intensity of their insight experiences.

Measures. Since every participant solved the same two problems, the difference in the degree of restructuring of the two problems was constant. Problem # 10 was perceived to be significantly more difficult than problem # 4 ($p < 0.001$) in Experiment 1. The difference in the computed change score of these two problems is also substantial: 12 units before the adjustment for problem #4, and 7 units taking into account the adjustment.

Solution time was used as an objective measure for problem difficulty. The solution time was measured from the moment the folder containing the problem was opened to the moment the participant made the correct move.

The amount of tension released after the solution moment for each of the problems was measured as the height of the tension curve drop after the solution point. After solving each problem, participants were asked to draw a curve that represented the changes in their feeling of tension while they were working on the problem and after they found the solution. Tension was defined as a feeling of pressure, uneasiness, anxiety, or stress they might have experienced.

Participants were given a letter-size graph with two axes. The vertical axis was marked “Tension/Stress.” The participants were told that the point of intersection with the horizontal axis is the “zero” level of tension, and that intensity of tension increased with the increase in the height of the vertical axis. The horizontal axis was marked “Start” at the intersection with the vertical axis; and the “Solution” axis in the middle. The vertical axis was 13 cm long, and the horizontal axis was 20 cm long, with the “Solution” point located 10.5 cm from the “Start.” A blank Tension Graph is presented in Appendix F. Participants were told that the horizontal axis represented their progress on the problem. They were instructed to show any changes (increases/decreases/stability) in their feeling of tension throughout the whole process of working on the problem, including after the solution was found until the present moment in time. For the first problem, participants drew their tension curve relative to the zero level. For the second problem, they drew their tension curve relative to the first problem on the same graph to show any differences in their experience of tension between the two problems. An example of a completed tension graph is shown in Appendix G.

The intensity of insight for each problem was measured as the distance from the beginning of the insight intensity scale to the point marked by a participant corresponding to this problem.

After solving both problems, participants were asked to show the difference in the intensity of their insight moment for the two problems they solved on a scale. The scale had a label “Weak insight” on the extreme left of the continuum, an arrowhead pointing to the right, and a label “Intense insight” on the extreme right. The scale was 14 cm long. Other than these two labels and the arrowhead, there were no other markings. The Insight Scale in presented in Figure 4.

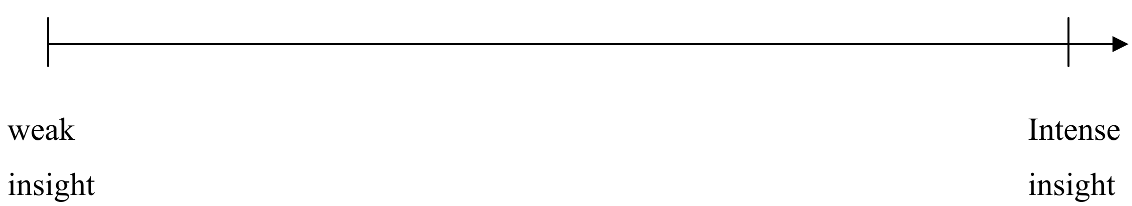


Figure 4: Insight Intensity Scale

The participants were asked to position two points on the scale representing the two solution moments of the two problems they solved. The instruction was to position the two points such that the distance between the points relative to each point’s distance to the beginning of the scale represented the relative difference in the intensity of their insight experience for the two problems.

3.3.2. Results

The purpose of this experiment was to investigate whether there will be a difference in the experience of insight for individuals who solve two problems with different degrees of difficulty under the same conditions.

Only 28 out of 42 participants solved both problems during the experimental session. All participants were able to solve their first problem, but 14 people could not solve the second problem. Only the data from the 28 participants who solved both problems during the experiment was analyzed since the interest of this study was the comparison of individual's experience of insight for the two problems.

The two problems for this experiment were selected such that one of them, problem # 10, required a greater degree of restructuring and, as a result, was significantly more difficult than the other problem, problem #4. Consequently, the following experimental predictions were formulated and tested:

1. Problem # 10 will take longer to solve than problem # 4;
2. Problem # 10 will be associated with higher levels of tension released after the solution moment;
3. Problem # 10 will be associated with higher intensity of insight than problem # 4

Since the conditions of normality were not met for any of the measures, non-parametric tests were carried out. Table 10 presents medians and lower and upper quartiles for solution times in seconds, height of the tension curves' drop in cm after the solution point, and the intensity of insight scale measure in cm for the two problems.

It can be seen from Table 10 that problem # 10 has a much higher median solution time, as well as a higher median tension release level and a higher median intensity of insight level, than problem # 4. Single Wilcoxon Signed Ranks tests were computed to assess the statistical significance of the observed differences. Problem # 10 was found to be significantly higher on all three measures than problem # 4.

The corresponding test results were $Z = -3.92$, $p < 0.001$ for the solution time; $Z = -2.43$, $p < 0.01$ for the tension level; and $Z = -2.50$, $p < 0.01$ for the intensity of insight measure.

Table 10: Descriptive statistics for solution time (*seconds*), tension release (*cm*), intensity of insight (*cm*), *N* = 28

Variable	Lower *Q*	*Mdn*	Upper *Q*
Solution time, Problem # 4	24	62	116
Solution time, Problem # 10	158	390	932
Tension, Problem # 4	2	4	9
Tension, Problem # 10	4	6	10
Intensity of insight, Problem # 4	3	5	9
Intensity of insight, Problem # 10	5	9	11

3.3.3. Discussion

The participants in this study were put in a situation where they actually had to solve two problems during the study. First, they solved an easier problem and then a more difficult one. The analysis of the solution times showed that the second problem was, in fact, significantly more difficult than the first. Both problems were from the same domain. Learning naturally occurs in the same domain from one problem to the next, and people generally get more comfortable with a problem if they have solved a similar one before. However, problem # 10 was so much more difficult that whatever learning had occurred during the solution of problem # 4 did not reduce the difficulty of problem # 10 by any substantial level.

The problem requiring a greater degree of restructuring for its solution was associated with a higher intensity of insight experience. However, in this case a more intense experience of insight can not be purely attributed to the degree of restructuring involved in a problem. The more difficult problem in this study was also associated with a higher level of tension generation and tension release after the solution as reported by participants.

Restructuring in a problem solving situation inevitably results in a certain level of tension in one’s life space. In insight problem solving, where the solution to a problem is achieved in a single moment of realization, the level of tension seems to be in a positive relationship with the degree of restructuring required to solve the problem. It is so in insight problem solving because up until that moment of realization, the moment of insight, there is no significant psychological locomotion towards the solution. A greater degree of restructuring might naturally take longer to see, thereby keeping the solver in an undesirable region which does not have any significant outlet for one's tension. The interaction of various forces creates tension in the life space, which in problem solving situations can only be completely released through reaching the solution region. In such situations, tension increases until the solution is found.

This study does not provide a direct test for Hypothesis 1 because the level of tension increased along with the increased degree of restructuring. However, one could argue that since the environment for the two problems was kept constant, the increase in tension level and the higher intensity of the insight experience reported for the second problem both resulted from the greater degree of restructuring. Thus, in this situation, a greater degree of restructuring led to a higher level of tension, and the combination of the two led to a more intense experience of insight.

It is impossible to separate the degree of restructuring and tension in a true problem solving situation. Tension always arises in a problem solving situation and changes along with the degree of restructuring. However, the degree of restructuring is not the only contributor to tension levels in problem solving. The following two experiments examine how a manipulation of tension affects the intensity of insight experience without changing the degree of restructuring.

3.4. Experiment 3

The purpose of Experiment 3 was to investigate the effect of tension on the intensity of the insight experience during the problem solving process while keeping the degree of restructuring constant thereby testing Hypothesis 2. Hypothesis 2 states that when the solution to the same problem is found with different levels of tension, the intensity of the insight experience will be greater where a greater amount of tension was released with the solution of the problem.

In this experiment, the level of tension generated during the problem solving process was manipulated in two different ways. First, an explicit time constraint on the duration of the problem solving process was imposed. The time constraint acted as an additional induced force toward the direction of the goal, which increased the overall resultant force in the field. The amplified resultant force led to an increase in the level of tension in the system through the interaction of the driving and restraining forces. According to the field theory, the level of tension depends heavily on the strength of the opposing forces in the field. The strength of the restraining force of the impassable barrier upon the encounter is equal to the strength of the driving force towards it. Whenever the driving force towards the barrier is increased it leads to the increase in the barrier's restraining force, and thus a higher level of tension.

The second method of tension manipulation in this experiment was achieved through changing the valences of the sub-regions in the PSR. This also led to a change in the interaction of the forces and the level of tension in the field. A reward was offered for the successful solution of the problem. The active participation of the problem solvers in the choice of a particular reward and the constant physical presence of the reward throughout the whole problem solving process ensured its 'psychological presence' for the participants. Moreover, the reward chosen by a participant was said to be taken back if the participant did not solve the problem. These conditions might have increased the positive valence of the goal sub-region, but on the other hand, they also increased the negative valence of the "Failure" sub-region since a negative (or socially undesirable) consequence was associated now with not solving the problem. The act of offering a reward in exchange for solving the problem also slightly strengthened the induced force on the problem solver coming from the expectation of the experimenter that the participant would solve the problem. This increase in the induced force was probably not as strong as the one caused by the time constraint. The greater positive valence of the goal sub-region increased the goal force while the greater negative valence of the "Failure" sub-region increased the repelling force from that region. Both of these changes, in addition to a slightly boosted induced force, contributed to the overall increase in the resultant force towards the goal, and to a higher level of total tension.

To make the detection of the changes in the tension levels possible, the participants in this study first solved a "benchmark" problem followed by an experimental problem associated with the tension manipulation. After solving the experimental problem, the participants compared their experience of the experimental problem to the benchmark problem. The reason the participants were asked to solve two problems was to allow for a meaningful comparison of their experience in two similar and proximal situations. The study employed a between-participant experimental design with three groups: the control group and one treatment group for each tension manipulation approach.

3.4.1. Method

Participants. 120 undergraduate students from the University of Waterloo participated in this study in exchange for a partial course credit. Three of the 120 participants could not solve one of the two problems during the experimental session and were excluded from the analysis.

Materials. The ideal stimuli for this study would have been two problems with the same degree of difficulty from the same problem domain. However, in matchstick equations, problems with exactly same degree of restructuring are also similar in their solution approaches. This could be detrimental to a study where participants have to solve two problems. As was shown by Knoblich et al. (1999), once the constraint associated with the difficulty of a problem is relaxed for the first time, it stays relaxed, considerably decreasing the experienced difficulty on the second problem that involves the same solution approach.

The two stimuli problems for this study were chosen such that they involved different approaches to the solution, thereby minimising the learning effect from the first problem to the second. The stimuli for this study were the two matchstick problems – problem # 4 and problem # 7 – from Table 4.

- Benchmark problem: IX = III + I, solution IV = III + I (problem # 4, Table 4)
- Experimental problem: IV = III – I, solution IV – III = I (problem # 7, Table 4)

These two problems were compared to each other in pair 11 of Experiment 1, and the results indicated that problem # 4 was judged to be significantly more difficult than problem # 7 ($Z = -2.04$, $p < 0.05$, $N = 37$).

Since problem #4 was perceived to be more difficult in Experiment 1, it was used as the benchmark problem and was solved first in this study. Tension manipulations were applied to the second problem, problem #7, which was considered to be the easier of the two, thereby allowing for a stronger test of the hypothesis.

The problems were set up using brown coffee sticks for representation inside two taped together legal-size yellow folders. The sticks were placed on one side of the folder while the other side of the folder was used to cover the problem. When the folder was closed, the participant could not see the problem. The folder was located on the table in front of the participants. To start working on the problem, a participant opened the folder to reveal the puzzle. This allowed for a clear marking of the beginning of the problem solving process.

Procedure. The study was conducted in a laboratory setting with one participant at a time. Each experimental session lasted one hour and was videotaped. After arriving at the laboratory and signing the consent forms, the participants received videotaped instructions about the study presented to them on a computer screen. The transcript of these instructions is included in Appendix H. In the video instructions, participants were told that the study investigated people's experience when they solve problems and that they will be asked to solve two problems and answer several questions about their experience after they have solved the problem. Insight was defined as an “Aha! I see now!” experience; as an experience of a "light bulb flashing" above one’s head (see Appendix H for more details).

The participants were then taken through the experimental procedure in a training session. All activities, measures and sequences of events were explained during this training. During this training, the participants were asked to solve two 3-letter-word anagrams in place of the matchstick problems. The purpose of this training was to ensure that the participants were comfortable with the experimental procedure and the laboratory setting. After this training, the participants received additional video recorded instructions on matchstick equation problems and matchstick representation of the symbols was presented on a computer screen. An example of a matchstick equation along with its solution was also shown on the screen (problem: II + II = II, solution: II + I = III). Each acceptable symbol (Roman numerals from 1 to 12 and five operations) was constructed with the coffee sticks and presented on the screen along with its meaning. A list containing Roman numerals from 1 to 12 and five mathematical operations was available to the participants at all times during this study. This list is presented in Appendix B. The transcript of these instructions is provided in Appendix H. While viewing the video instructions, the participants had an opportunity to pause the recording if they had any questions. While the participants were watching the second set of video instructions, the benchmark problem was constructed inside the folder on the table such that the participants did not see it. After the instructions were over, the participants were instructed to start working on their first problem. They could start any time by simply opening the folder and revealing the problem. There was no time constraint for the first problem.

After a participant declared that they had solved the problem, the solution was verified and the participant was asked a series of questions. First, each participant was asked to draw their tension curve for the first problem. To avoid making an assumption regarding participants’ experience of insight, the participants were asked to report whether they had an experience of insight when they solved the problem (e.g. Bowden et al., 2005). If the answer was affirmative, the participant was asked to go back in their mind to the moment when they realized the solution to the problem and “get a good feel for that moment, especially its intensity.” The procedure for the first problem was the same for all three experimental groups.

The experimental problem then was set up inside the folder. Before participants started working on the second problem, they were asked to remember their solution moment on the first problem one more time. The instructions given to the participants immediately before their second problem differed among the experimental groups.

- The control group received the same instructions as for the first problem (i.e. no time constraint);
- Treatment group 1, the time pressure group, were instructed that they had only three minutes to work on the second problem, and that a timer with a ticking noise will be used to remind them of the time limit;
- Treatment group 2, the chocolate group, were asked to choose a chocolate bar that they would have liked to receive if they solved the second problem. A choice of four kinds of chocolates was given to them. The chosen bar was placed on the table next to the folder containing the problem, and the participants were told that the chocolate bar would be taken back if they did not solve the second problem.

The participants could start working on their second problem at any time by opening the folder and revealing the puzzle. In all experimental groups, including the time pressure group, the participants were allowed to work on the second problem for as long as they wished, and they were stopped only if they could not solve the problem within 20 minutes. A metronome was used for the time pressure group to generate the ticking sound in a B-flat tone with one second intervals for the whole duration of the problem solving. The metronome was turned on as soon as the participant opened the folder, and turned off as soon as the participant declared that they found the solution. For the chocolate group, the chocolate remained on the table in front of the participants for the whole duration of the problem solving. After the participants solved the problem, they were told that they could keep the chocolate.

After a participant solved the second problem, their solution was verified and they were asked a series of questions. First, they were asked to draw their tension curve relative to that of the first problem. The participants drew both of their tension curves on the same sheet using different coloured markers. The participants were instructed to show any differences in their feeling of tension on the second problem compared to the first in their drawing, using the distance between the curves to illustrate the proportionality of the difference. Participants were then asked to show the difference in the intensity of their insight moment between the problems on an unmarked scale (Figure 4). They were asked to position two points on the scale, representing the two solution moments, and position them such that the relative distance between the points and the beginning of the scale represented the difference in the intensity of their insight experiences.

The last task for the participants was to provide their subjective judgment of the relative difficulty for the two problems on a 10-point scale. The participants were to mark both problems on the same scale showing the difference in their own experience of difficulty by distancing the points accordingly. The ends of the scale were marked as "Very easy" for 0 and "very difficult" for 10 (see Figure 5).

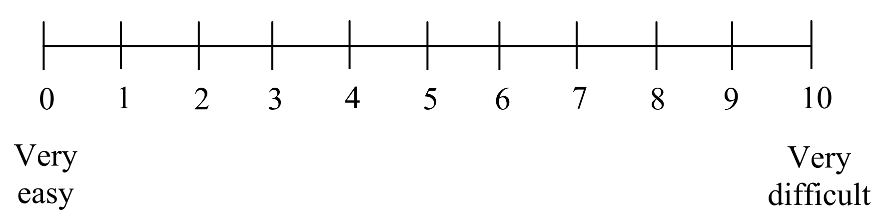


Figure 5: Problem difficulty scale

Participants were assigned randomly to one of the three experimental groups. There were 40 participants in both the control and the chocolate group, and there were 37 participants in the time pressure group.

Measures. Since every participant solved the exact same two problems, the **objective difference in the degree of restructuring** of the two problems was constant for all the experimental groups.

Solution time was measured in seconds from the moment a folder with the problem was opened to the moment the participant made the correct move. The **relative solution time** was of interest in this study as a relative measure of the objective difficulty and performance. The relative solution time was calculated as the difference between the solution times on the experimental and the benchmark problems. It was calculated by subtracting the solution time for the benchmark problem from the solution time for the experimental problem.

After solving each problem, participants were asked to draw a curve that represented the changes in their feeling of tension while they were working on the problem and after they found the solution. For the first problem, participants drew their tension curve relative to the zero level. For the second problem, they drew their tension curve relative to the first problem on the same graph to show any differences in their experience of tension between the two problems. Tension was defined as feeling of uneasiness, anxiety, stress or discomfort they might have experienced. The graph used was the same as in Experiment 2, and it is shown in Appendix F. The participants received detailed instructions about this graph during the training process.

The **relative amount of tension release** for the two problems was of interest in this study. The relative amount of tension release was calculated as the difference in the decreases of tension between the experimental and the benchmark problems after the solution point. It was computed by subtracting the height of the tension curve drop after the solution for the benchmark problem from that of the experimental problem.

The participants indicated the relative intensity of their insight experience on the two problems on a scale. The distance from the beginning of the insight intensity scale to each of the two points marked by a participant was measured in centimetres. The **relative intensity of insight** of the two problems was of interest in this study. The relative intensity of insight was calculated as the distance between the two points on the insight intensity scale by subtracting the insight measure for the benchmark problem from that for the experimental problem.

The participants indicated the relative difficulty of the two problems on a 10-point scale. The **relative perceived difficulty** of the two problems was of interest for this study. The relative perceived difficulty was calculated by subtracting the difficulty measure for the benchmark problem from that for the experimental problem.

The control group was compared to the two treatment groups on four measures:

- Relative solution time;
- Relative amount of tension release;
- Relative intensity of insight;
- Relative perceived difficulty.

It is worth noting that all of the measures could assume both positive and negative values. A positive value on any of the measures would indicate that the experimental problem scored higher than the benchmark problem, and a negative value would indicate that the experimental problem scored lower.

3.4.2. Results

Experiment 3 was designed to test the effect of tension on the intensity of the insight experience while keeping the degree of restructuring constant. The participants compared their experience on the two problems they solved in terms of level of tension, the intensity of their ‘aha!’ moment, and the perceived difficulty. The relative measures reported in participants’ experiences were compared between the control group and the two treatment groups. The following experimental predictions from Hypothesis 2 were tested:

- The relative amount of tension release will be greater for the two treatment groups than for the control group;
- The relative intensity of the insight experience will be greater for the two treatment groups than for the control group;
- The relative perceived difficulty will be greater for the two treatment groups than for the control group;
- The relative solution times (as objective measures of difficulty) will be higher for the two treatment groups than for the control group;
- There will be a positive correlation between the relative amount of tension release and the relative intensity of the insight experience;
- There will be a positive correlation between the relative amount of tension release and the relative perceived difficulty and the relative solution time.

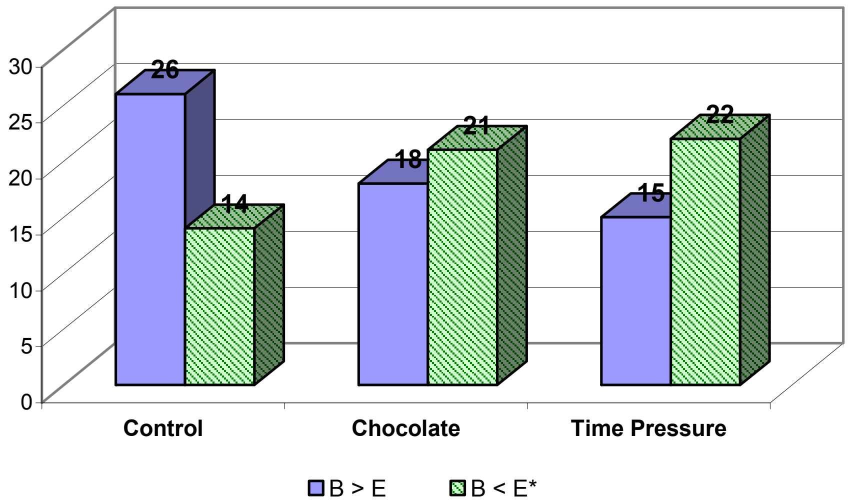
It is worth noting that all participants indicated that they had experienced insight when they solved the two matchstick problems in this experiment.

3.4.2.1. Dichotomous representation

First, to examine an overall pattern in the results, the frequencies of positive and negative values for each of the measures were counted across the three groups. A positive value on any of the measures indicated that the experimental problem scored higher than the benchmark problem, and a negative value indicated that the experimental problem scored lower.

The binary frequencies of the solution times are presented in Figure 6 for the three groups.

In the control group, the majority of the participants spent longer solving the benchmark problem than the experimental one. In the two treatment groups, the chocolate and the time pressure, the majority of the participants spent longer solving the experimental problem. The observed differences in the frequencies were significant ($\chi^2 = 5.2$, $df = 2$, $N = 116$, $p < 0.05$ 1-tailed).

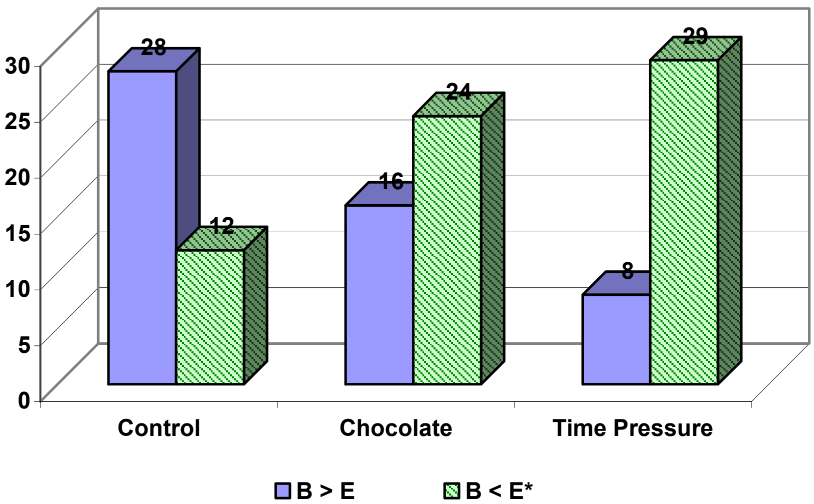


* B > E: the benchmark problem took longer than the experimental problem; B < E: the experimental problem took longer than the benchmark

Figure 6: Binary report of the solution times in the three groups

The binary frequencies of the tension release are presented in Figure 7 for the three groups.

In the control group, the majority of the participants reported higher levels of tension release on the benchmark problem than on the experimental one. In the two treatment groups, the chocolate and the time pressure, the majority of the participants reported higher levels of tension release on the experimental problem. The observed frequencies were significantly different ($\chi^2 = 18.7$, $df = 2$, $N = 117$, $p < 0.001$).

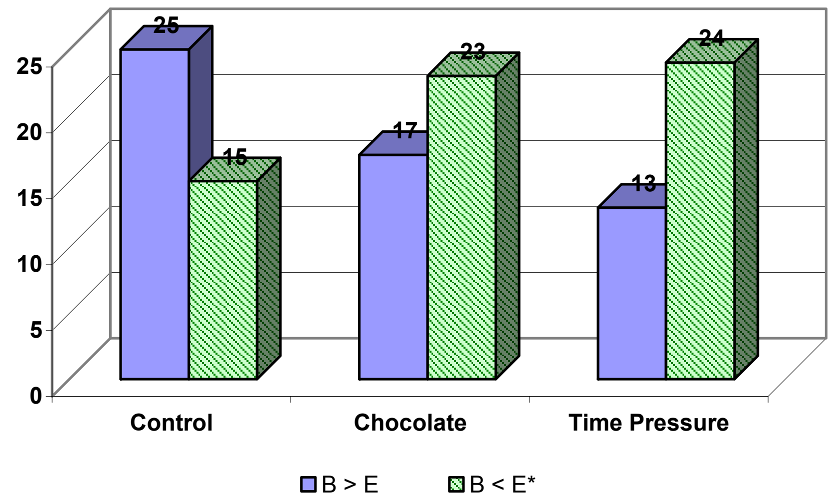


* B > E: More tension was released on the benchmark problem than on the experimental problem; B < E: More tension was released on the experimental problem than on the benchmark

Figure 7: Binary report of the amount of tension released after the solution in the three groups

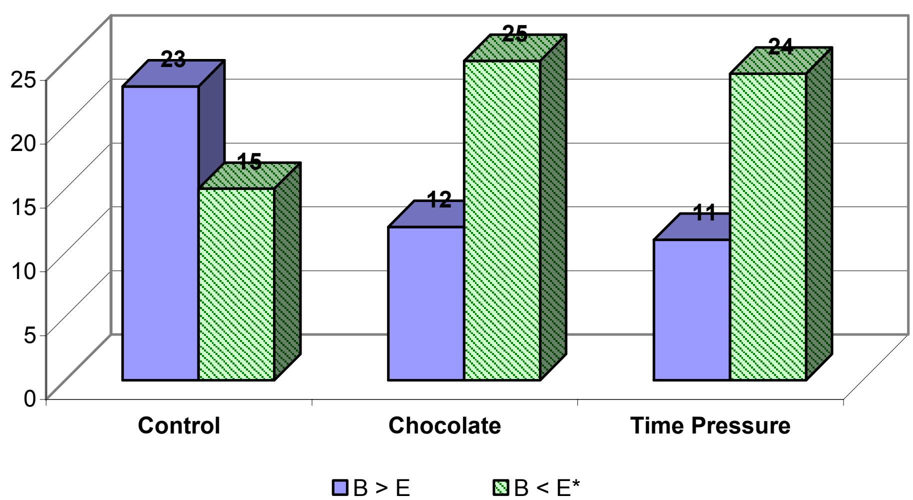
The binary frequencies of the intensity of the insight experience are presented in Figure 8 for the three groups. In the control group, the majority of the participants reported more intense experience of insight on the benchmark problem than on the experimental one. In the two treatment groups, the chocolate and the time pressure, the majority of the participants reported more intense experience of insight on the experimental problem. The observed frequencies were significantly different ($\chi^2 = 6.3$, $df = 2$, $N = 117$, $p < 0.05$).

The binary frequencies of the judgment of difficulty are presented in Figure 9 for the three groups. In the control group, the majority of the participants perceived the benchmark problem to be more difficult. In the two treatment groups, the chocolate and the time pressure, the majority of the participants perceived the experimental problem to be more difficult. The observed difference was significant ($\chi^2 = 8.4$, $df = 2$, $N = 110$, $p < 0.05$).



* B > E: Insight was reported to be more intense on the benchmark problem than on the experimental problem; B < E: Insight was reported to be more intense on the experimental problem than on the benchmark

Figure 8: Binary report of the intensity of the insight experience in the three groups



* B > E: The benchmark problem was more difficult than the experimental problem; B < E: The experimental problem was more difficult than the benchmark

Figure 9: Binary report of the difficulty judgment in the three groups

The above report illustrates that while the benchmark problem scored higher on all of the measures in the control group for the majority of the participants, this switched in the two experimental groups such that the experimental problem scored higher in the majority of the cases. To investigate the observed differences, further analysis of the relative measures was performed.

3.4.2.2. Analysis of the relative measures

None of the four measures satisfied the conditions of normality and homogeneity of variance required for the parametric statistical analysis.

To correct for this problem, all measures were normalized within each participant by dividing the obtained value of the relative measure by the range used by that participant for that particular metric. This also reduced the effects of individual differences in using the graphs and scales. The following formula was used for normalising

$$RM_{ijN} = \frac{M_{ij}(Ep) - M_{ij}(Bp)}{M_{ij}(Ep) + M_{ij}(Bp)}$$

RM_{ijN} is a normalised relative value of the measure i (i = 1 to 4) for participant j (j = 1 to 117);
$M_{ij}(Ep)$ is a value of the measure i for participant j on the experimental problem;
$M_{ij}(Bp)$ is a value of the measure i for participant j on the benchmark problem.

This procedure emphasised the proportionality of the difference between the two problems.

After this procedure was administered to the data, all the measures were distributed in the range (-1; 1). Computation of the normality and homogeneity of variance tests to the normalised measures showed that three of the measures (the relative amount of tension release, the relative intensity of insight, and the relative perceived difficulty) satisfied the requirements of normality and equality of variance. However, the relative solution time measure did not. In the subsequent analysis normalised scores were used for the relative tension, insight, and the perceived difficulty. The relative solution times were analysed separately with ordinal statistical procedures.

First, the relative solution times as objective measures of difficulty were analyzed. The medians and quartiles of the relative solution times are reported in Table 11 for the three experimental groups.

The observed differences in the relative solution times between the control group and the two treatment groups were noteworthy based on the results of the Kruscal-Wallis test ($\chi^2 = 5.9$, $df = 2$, $N = 117$, $p < 0.05$). Further, single U tests showed that the relative solution times were significantly higher in the time pressure group than in the control group ($U = 512.5$, $Z = -2.32$, $p < 0.05$, 1-tailed), and significantly higher in the chocolate group than in the control group ($U = 607$, $Z = -1.86$, $p < 0.05$, 1-tailed).

Table 11: Relative Solution Times (in seconds) in Experiment 3

Statistic	Experimental group		
	Control Group	Chocolate group	Time pressure group
Lower *Q*	- 160	- 74	- 44
Mdn	- 40	9	10
Upper *Q*	33	200	102

The differences in the relative solution times between the control group and the treatment groups indicated that the participants in the two treatment groups took significantly more time to solve the experimental problem relative to their own performance on the benchmark problem than the participants in the control group.

The means and standard deviations for the three dependent variables – the relative amount of tension release, the relative intensity of insight, and the relative perceived difficulty – are reported in Table 12 for the three experimental groups.

Table 12: Means and standard deviations of the relative measures of tension, insight, and perceived difficulty

Group	*N*	Dependent Variable					
		Relative amount of tension release		Relative intensity of insight		Relative perceived difficulty	
		mean	SD	mean	SD	mean	SD
Control	40	- 0.20	0.35	- 0.10	0.46	- 0.08	0.45
Time pressure	37	0.20	0.31	0.13	0.45	0.18	0.44
Chocolate	40	0.02	0.39	0.11	0.44	0.16	0.47

The multivariate ANOVA was computed to assess the effect of the experimental condition (the independent factor) on the relative amount of tension release, intensity of insight, and perceived difficulty (the dependent variables). The results showed significant differences among the groups (*Pillai's trace* test $F = 4.44$, $p < 0.001$). The subsequent univariate analyses revealed significant differences among the groups in the relative amount of tension release ($F(2, 114) = 11.79$, $p < 0.001$). Dunnett's post hoc comparisons revealed that the relative amount of tension release was significantly higher in the time pressure group than in the control group ($p < 0.001$, 1-tailed) and significantly higher in the chocolate group than in the control group ($p < 0.01$, 1-tailed).

There was also a significant effect of the experimental tension condition on the relative intensity of the insight experience ($F(2, 114) = 3.66$, $p < 0.05$). Dunnett's post hoc comparisons revealed that the relative intensity of insight was significantly higher in the time pressure group than in the control group ($p < 0.05$, 1-tailed) and significantly higher in the chocolate group than in the control group ($p < 0.05$, 1-tailed).

The experimental condition had a significant effect on the relative perceived difficulty ($F(2, 114) = 4.16$, $p < 0.05$). Dunnett's post hoc comparisons revealed that the relative perceived difficulty was significantly higher in the time pressure group than in the control group ($p < 0.05$, 1-tailed) and significantly higher in the chocolate group than in the control group ($p < 0.05$, 1-tailed).

Spearman correlation coefficients were computed among the pairs of the relative measures (non-normalized). A highly significant correlation was observed between the pairs of all the dependent variables. Spearman correlation coefficients between the pairs of dependent variables are reported in Table 13.

Table 13: Spearman correlation coefficients between the relative measures (N = 117)

Variable (relative measure)	Variable (relative measure)		
	Amount of tension release	Intensity of insight	Perceived difficulty
Amount of tension release	.	.	.
Intensity of insight	0.733**	.	.
Perceived difficulty	0.790**	0.812**	.
Solution times	0.728**	0.725**	0.846**

** $p < 0.001$

3.4.2.3. Discussion

This experiment was designed to examine the effect of changing the level of tension during problem solving on the intensity of the insight experience. The difference in the degree of restructuring was kept constant in all three experimental groups since all participants solved the same two problems. The level of tension generated during the solution process of the experimental problem was manipulated by imposing a time constraint in one group and by adjusting the valences of the goal and the failure sub-regions through a reward in the other group. The third group, the control, solved the problem without any experimental manipulation of tension. The participants in all groups compared their experience of the experimental problem to another problem from the same domain they solved immediately before the experimental problem. The relative levels and changes in tension, intensity of the insight experience and perceived difficulty were reported by the participants. The relative differences in the participants' experience were compared between the control group and each of the treatment groups.

In the time pressure group, an imposed time constraint with a constant audio signal created an additional induced force in the direction towards the goal in the life space of the participants. Besides creating an induced force, the time constraint could have also increased the valence of the goal sub-region as the problem might have seemed more difficult under the time constraint, and thus more attractive to solve. Adding an induced force and possibly increasing the goal valence produced a stronger resultant force towards the goal. A barrier poses a greater resistance when it is encountered with a greater force ultimately creating a higher level of tension. Although not all participants responded equally to the time constraint, the overall effect showed a significant increase in tension.

In the chocolate group, participants were asked to choose one of four kinds of chocolate bars. They were told that they will receive the chocolate if they solve the problem, and the chocolate was placed on the table where participants were seated directly next to the problem they needed to solve. Also, the participants were told that the chocolate will be taken away if they did not solve the problem. It is unlikely that a $1 chocolate bar has any substantial value for an average Canadian undergraduate student. It was not the monetary value of the item itself that had an effect on the participants, but rather it was their active participation in the selection of the item, its constant presence during the problem solving process, and the social consequences associated with not solving the problem (i.e. the chocolate would be removed if the problem was not solved). These factors increased the negative valence of the failure sub-region by emphasising the possibility of not solving the problem which might have been socially undesirable for the participants.

It is worth pointing out that there was no reward offered for the solution of the benchmark problem in the study. The reward was offered only for the solution of the experimental problem which was the second problem participants solved. Just the act of offering a reward itself, regardless of its value, increased the induced force on the participant to act in the direction toward the goal. Also, the valence of the goal sub-region might have increased not mainly because of the value of the chocolate, but because somebody else (i.e. the experimenter) placed a greater value on solving this problem by offering a reward.

The experimental manipulation of tension was proven to be successful as the relative levels of tension reported by the participants in the two treatment groups were significantly higher than in the control group. The results of this experiment demonstrated that the level of tension in the life space during problem solving did not entirely depend on the degree of restructuring of the problem, and it could have been altered independently by changing the problem solving environment.

The experimental manipulation of tension had a significant effect not only on the relative perceived difficulty of the experimental problem, but also on the more objective measure of difficulty: the relative solution time. This finding could possibly be explained by the fact that the strength of the forces acting in the field was increased by the experimental manipulation which caused a higher level of tension.

The increased solution times due to the tension manipulation indicated that accumulating levels of tension in the life space might have affected the structure of the life space itself. This may have made the search for the correct solution more difficult. One possible explanation is that as tension had increased it might have spread in the problem solving region distorting its differentiation and leading to a "narrowing-down of the psychologically existing area," or even, possibly, causing a certain degree of "regression" (Lewin, 1935, 1951). In other words, higher levels of tension might affect the process of finding transformations in the search space and applying them to the problem.

There was a significant correlation between the relative amount of tension release, the relative perceived difficulty, and the relative solution times indicating that the level of tension did contribute to the experienced difficulty of the problem beyond the degree of restructuring required to solve it.

More importantly, the results demonstrated that the intensity of the insight experience was significantly higher in the two treatment groups that experienced an overall higher level of tension. All three groups in Experiment 3 solved the same two problems in the same order. The reported differences in the participants' experience could only be attributed to the experimental manipulation of tension. These results indicate that the experience of the intensity of insight depends not only on the problem itself, but on the dynamics of the solution process. The observed significant correlation between the relative intensity of the insight experience and the relative amount of tension release after the solution provide further support for this explanation.

The results of Experiment 3 provided considerable support for Hypothesis 2. Hypothesis 2 states that when the solution to the same problem is found with different levels of tension, the intensity of the insight experience will be greater in the case when a higher amount of tension is released with the solution to the problem.

3.4.3. Tension curve analysis: Results and discussion

The tension curves of the participants were analysed in greater detail to obtain more information on the dynamics of tension change during insight problem solving. It was proposed in the theoretical argument that tension in insight problem solving is either increasing or is sustained until the problem is solved; and at the moment of solution the tension from the problem solving region is rapidly released. Furthermore, it was also argued that no gradual tension reduction mechanism is available before the solution is found. The changes in the tension curves before the solution and after the solution were analysed to examine these theoretical propositions.

3.4.3.1. Overall changes in tension before and after the solution

First, in each tension curve drawn by the participants in Experiment 3, the total increase and total decrease were measured in centimetres before the solution point and after the solution point. Since these values were highly skewed, medians and quartiles are reported in Table 14.

For both problems, tension graphs exhibited a general pattern of a substantial increase before the solution and a major decrease after the solution. There was very little, if any, decrease in the tension curves before the solution point. Indeed, there was a significantly higher level of increase, rather than decrease, in the tension curves prior to the solution point for both problems across all groups (all Wilcoxon signed ranks tests *p*s < 0.001). There was also significantly more decrease in tension after the solution point than before the solution point for both problems across all groups (all Wilcoxon signed ranks tests *p*s < 0.001), and there were no visible increases in the tension curves after the solution point.

Table 14: Changes in the tension curves (centimetres)

			Experimental group					
			Control (*N* = 40)		Chocolate (*N* = 40)		Time Pressure (*N* = 37)	
			Problem		Problem		Problem	
			Benchmark	Experimental	Benchmark	Experimental	Benchmark	Experimental
Before the solution point	Increase in tension curve	Low *Q*	2.5	0.8	1.3	1.4	1.8	3.1
		Mdn	4.8	1.8	3.8	3.8	4.8	7.0
		Upper *Q*	6.9	4.5	5.8	7.3	6.1	9.1
	Decrease in tension curve	Low *Q*	0.0	0.0	0.0	0.0	0.0	0.0
		Mdn	0.3	0.0	0.0	0.0	0.0	0.0
		Upper *Q*	1.7	1.0	0.3	0.2	0.5	0.0
After the solution point	Increase in tension curve	Low *Q*	0.0	0.0	0.0	0.0	0.0	0.0
		Mdn	0.0	0.0	0.0	0.0	0.0	0.0
		Upper *Q*	0.0	0.0	0.0	0.0	0.0	0.0
	Decrease in tension curve	Low *Q*	3.6	1.6	3.3	3.2	2.1	6.2
		Mdn	5.6	2.9	5.5	6.2	6.0	7.5
		Upper *Q*	7.9	7.2	9.0	9.5	7.8	10.9

These facts indicated that the tension levels during the insight problem solving process in Experiment 3 had a tendency of building up without any significant decrease before the solution was found. This observation supports the field theory explanation of the dynamics of tension in insight problem solving. It was argued that due to the nature of insight problems, gradual progress towards the goal is not possible, and therefore no substantial outlets for tension release are available before the solution. Consequently, tension in the life space will sustain and increase until the solution is found. The above analysis of the tension curve results supports this argument.

Computed Kruskal Wallis tests revealed that there were no significant differences among the groups in any of the changes of tension for the *benchmark problem* (all *p*s > 0.3). However, there were significant differences in the total increase in tension before the solution and the decrease in tension after the solution for the *experimental (second) problem* across the groups. Subsequent *U* tests revealed that for the *experimental problem*:

Increase in tension curve before the solution point (Kruskal Wallis $\chi^2 = 8.6$, $df = 2$, $N = 117$, $p < 0.05$):

- [time pressure group] > [control group], $U = 465.5$, $Z = -2.8$, $p < 0.01$
- [chocolate group] > [control group], $U = 649$, $Z = -1.5$, $p = 0.07$

Decrease in tension curve after the solution point (Kruskal Wallis $\chi^2 = 14.3$, $df = 2$, $N = 117$, $p < 0.001$):

- [time pressure group] > [control group], $U = 388.5$, $Z = -3.6$, $p < 0.001$
- [chocolate group] > [control group], $U = 561.5$, $Z = -2.3$, $p < 0.01$

Significantly more tension was released in each of the treatment groups than in the control group after the experimental problem was solved, which further supports the assertion that tension manipulation was successful in the two treatment groups.

3.4.3.2. Tension relationship between the benchmark and the experimental problems

The relationship between the two problems, the benchmark problem and the experimental problem, in terms of tension curve increases before the solution point and decreases after it are presented in Table 15 for each group. These relationships were assessed based on single Wilcoxon signed ranks tests.

Table 15: Tension curve relationships between the benchmark (B) and the experimental (E) problems

Tension curve change	Experimental group		
	Control	Chocolate	Time pressure
Increase in tension curve before solution	B > E*	B < E (p = 0.09)	B < E*
Decrease in tension after the solution	B > E*	B < E (p = 0.24)	B < E*

* $p < 0.05$

In the control group, a greater tension increase before the solution and a greater decrease after the solution were reported for the benchmark problem than for the experimental problem. As can be seen from Table 15, this difference between the problems was eliminated in the chocolate group, but the reverse effect did not reach a significant level. In the time pressure group, however, the difference in tension curve changes between the problems was a significant reversal compared to the control group. This indicates that the tension manipulation did not produce the same effects in the two treatment groups. The tension manipulation in the chocolate group considerably reduced the difference in the increase and decrease of tension between the two problems – significance of the higher tension for the benchmark problem found in the control group had disappeared in the chocolate group. The tension manipulation in the time pressure group resulted in significantly higher tension for the experimental problem, which reversed the difference in the amount of tension release observed in the control group.

3.4.3.3. Pattern of tension increase

For a further, more detailed analysis, each tension curve was decomposed into three equal parts before the solution point representing the beginning, middle and end of the solution process. Since the decreases in the tension curves were not significant, they were not included in the subsequent analysis. Increases in the tension curves in each of these parts were measured in centimetres and are reported in Table 16.

Table 16: Increases in tension curves (centimetres)

Group	Statistic	Problem					
		Benchmark problem			Experimental problem		
		Beginning	Middle	End	Beginning	Middle	End
Control	Low *Q*	0.6	0.5	0.0	0.2	0.0	0.0
	Mdn	1.6	1.4	0.6	0.8	0.5	0.0
	Upper *Q*	3.6	2.4	1.7	2.8	1.6	1.0
Chocolate	Low *Q*	0.2	0.1	0.0	0.2	0.1	0.0
	Mdn	1.5	0.9	0.3	0.9	1.2	0.7
	Upper *Q*	2.6	2.1	1.7	2.0	2.4	2.0
Time pressure	Low *Q*	0.4	0.1	0.1	0.4	0.8	0.2
	Mdn	1.2	0.9	1.0	1.7	1.8	1.5
	Upper *Q*	3.0	2.1	1.9	3.5	3.3	2.9

The differences in the increases in tension across these three different intervals of the curves were analyzed.

In the control group, for the benchmark problem, the biggest increase in the tension curves took place during the beginning and middle regions. Although, there was a greater increase in the beginning region compared to the middle, this difference was not found to be significant ($Z = -1.39$, $N = 40$, $p = 0.16$, 2-tailed for Wilcoxon signed ranks test). The increase in the tension curve at the end interval, however, was significantly less than in both the beginning and the middle ($Z = -3.00$, $N = 40$, $p < 0.01$, 2-tailed for the comparison between the end and the beginning; and $Z = - 2.88$, $N = 40$, $p < 0.01$, 2-tailed for the comparison between the end and the middle).

For the experimental problem in the control group, the same pattern of tension increase was repeated, with the beginning region having the highest increase in tension, followed by the middle region, and the least increase in tension taking place during the end interval of the curves. The associated Wilcoxon signed ranks test results are as follows: beginning compared to the middle: $Z = - 1.70$, $N = 40$, $p = 0.09$, 2-tailed; beginning compared to the end region: $Z = - 3.52$, $N = 40$, $p < 0.01$, 2-tailed; end compared to the middle region: $T = -3.01$, $N = 40$, $p < 0.01$, 2-tailed.

For the benchmark problem in the other two experimental groups, the pattern of tension increase throughout the regions of the curve was the same as in the control group. The tension increased the most in the first third of the curves, slowed down slightly (although, not significantly) in the middle section, and slowed down considerably in the last third of the graphs.

Kruskal Wallis tests were computed to assess the significance of the differences among the groups in tension increases at different stages of problem solving (i.e. beginning, middle and end on the benchmark problem). There were no significant differences among the groups in tension changes on this problem during all three intervals ($\chi^2 = 1.57$, $df = 2$, $N = 117$, $p = 0.46$ for the beginning interval; $\chi^2 = 1.77$, $df = 2$, $N = 117$, $p = 0.41$ for the middle interval; and $\chi^2 = 2.9$, $df = 2$, $N = 117$, $p = 0.24$ for the end interval).

The above analysis implies that the tension curve of the benchmark problem represented an increasing graph at a decreasing rate in all three groups.

However, for the experimental problem, the pattern of tension increase was different in the two treatment groups. In the chocolate group, there was a steady increase in tension throughout all three intervals of the curve with no significant differences among the intervals. The associated results of Wilcoxon signed ranks tests are as follows: the beginning interval compared to the middle: $Z = -0.66$, $N = 40$, $p = 0.51$, 2-tailed; the beginning interval compared to the end: $Z = -0.02$, $N = 40$, $p = 0.98$, 2-tailed; the end interval compared to the middle: $Z = -0.58$, $N = 40$, $p = 0.56$, 2-tailed.

In the time pressure group for the experimental problem, the tension increased the most in the beginning and the middle part with no significant difference between the two intervals ($Z = -0.02$, $N = 37$, $p = 0.98$, 2-tailed). The end interval showed significantly less increase than the middle ($Z = -3.02$, $N = 37$, $p < 0.01$, 2-tailed), but not significantly less than the beginning ($Z = -0.96$, $N = 37$, $p = 0.34$, 2-tailed).

Computed Kruskal Wallis tests revealed that there were significant differences among the groups in the level of tension increase for the experimental problem in the middle and the end sections of the curves ($\chi^2 = 1.73$, $df = 2$, $N = 117$, $p = 0.42$ for the beginning interval; $\chi^2 = 8.04$, $df = 2$, $N = 117$, $p < 0.05$ for the middle interval; and $\chi^2 = 14.26$, $df = 2$, $N = 117$, $p < 0.001$ for the end interval). The subsequent *U* tests showed that in both the chocolate group and the time pressure group there was significantly greater tension increase on the experimental problem in the end section of the curve than in the control group ($U = 520.5$, $Z = -2.8$, $p < 0.01$ for the chocolate group compared to the control; and $U = 396.0$, $Z = -3.6$, $p < 0.001$ for the time pressure group compared to the control). There was no significant difference in tension increase in the middle section of the curve between the control group and the chocolate group ($U = 669.0$, $Z = -2.0$, $p = 0.2$). However, there was significantly more increase in the middle section of the curve in the time pressure group than in the control group ($U = 467.0$, $Z = -2.8$, $p < 0.01$).

These results suggest that the 'usual' pattern of tension increase during insight problem solving in Experiment 3 had a shape of an increasing function at a decreasing rate, with the greatest rise in tension occurring during the beginning of the process, a lesser rise in the middle and the least rise during the end interval. This was observed for the benchmark problem in all three groups and for the experimental problem in the control group. That is, for all of the problems that were solved without any experimental manipulation of tension.

In the beginning and the middle intervals, tension kept increasing due to unsuccessful attempts to solve the problems leading to the interaction of the forces in the field that increased the level of tension. Tension might have increased until it reached an 'uncomfortable' or 'detrimental' level for an individual. When this happened, the individual might have employed tension reduction mechanisms through the adjustment of the valences in the problem solving region in order to stabilise the tension increase. For example, an individual might have decreased the valence of the goal sub-region ("this is just a little matchstick problem, I have solved much more difficult problems"), or reduced the valence of the failure sub-region ("so what if I don't solve it, it is not a big deal"). The adjustment of the valences affects the strength of the forces in the field. When the forces are decreased, the tension might not increase for a while. For some individuals this might have happened sooner in the process while for others it might have happened later in the process. There was some evidence in the verbal descriptions of the participants indicating that certain tension reduction mechanisms were employed. However, this issue still requires further investigation as the evidence was limited only to those participants who volunteered this information. The general pattern found in Experiment 3 suggested that in the third and final interval, the tension was increasing at a significantly slower rate than in the previous two intervals.

In the two treatment groups, however, the rate of tension increase for the experimental problem was different from the 'usual' pattern described above. More specifically, there was a significantly greater increase in tension during the end interval in the two treatment groups compared to the control group. This indicates that in the two treatment groups the increase in tension was either sustained throughout the whole problem solving process, or kept increasing at a higher rate in the latter intervals of the curve than in the normal pattern.

This pattern might be related to the fact that there was a difference in the force field in the treatment groups. Presumably, stronger forces were acting in the life space in the two treatment groups than in the control group. This suggests that the tension reduction mechanisms in the two treatment groups were, perhaps, not as successful for the experimental problem, and could not achieve the same reduction in the rate of tension increase towards the end of the process.

3.4.3.4. Dynamics of tension after the solution

All tension curves exhibited a dramatic decrease right after the solution point. 99 out of 117 participants drew tension curves reaching the zero level after the solution point, indicating a complete release of tension. The remaining 18 participants showed a significant decrease in tension after the solution without reaching the zero level. The tension curves showed a significantly greater decrease in tension after the solution than an increase before the solution in both problems across all groups (all Wilcoxon signed ranks tests' *p*s < 0.05, 2-tailed). This was possible because many participants indicated their starting tension level at the beginning of problem solving as being above the zero level. There was significantly more decrease in tension after the solution than before the solution for both problems and across all the experimental groups.

These results demonstrate that the solution of a problem in Experiment 3 led to a significant tension reduction in the participants' life spaces.

Along with drawing their tension graphs, the participants were also asked to estimate the amount of time it took for their tension to decrease after the solution. The frequencies of the reported times in seconds are shown in Figure 10 for the two problems.

Approximately 40% of the participants reported an immediate tension release (i.e. 0 seconds) as soon as the problem was solved for both problems. For about 80% of the participants their tension had decreased to the lowest level in three seconds or less.

The results clearly indicate that the time it took for tension to decrease after the solution was significantly lower than the time it took to build it up during the problem solving process. There was no case where the solution time of a problem was lower than the reported tension decrease time. This time frame of only a few seconds for a release of the considerable amount of tension that was accumulated during the process points to a much faster rate of post-solution tension decrease than the pre-solution tension increase. These results support the assertion that a rapid tension release is associated with the solution of an insight problem.

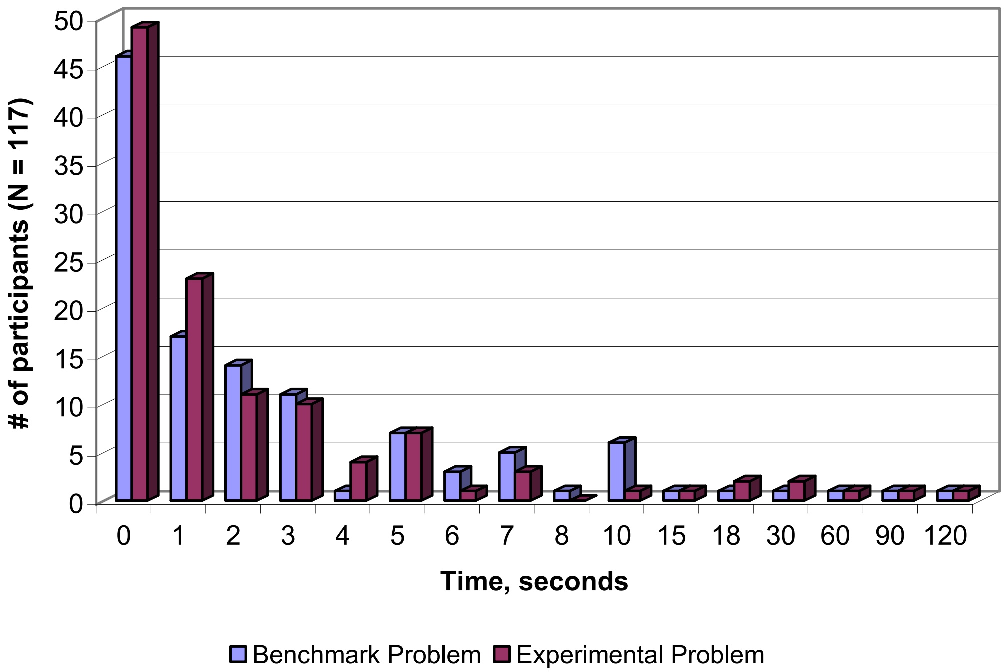


Figure 10: Frequencies of reported tension reduction times (in seconds) after the solution.

3.4.4. Discussion

This study achieved two purposes: first, it provided a direct test of Hypothesis 2 in a problem solving environment; and second, it allowed for an initial examination of the dynamics of tension during the insight problem solving process.

In this experiment the participants solved two insight problems from the same domain. The level of tension was manipulated in two out of the three groups for the second problem. The participants were asked to compare their problem-solving experiences in terms of the levels and dynamics of tension, the intensity of the insight experience and the perceived level of difficulty. The analysis of the tension curves drawn by the participants indicated that the tension manipulation techniques were, in fact, successful and led to higher relative levels of tension in the two treatment groups. The manipulation resulted in a stronger effect in the time pressure group than in the chocolate group.

There were no previously developed measures for the amount of tension release and the intensity of the insight experience. Both of these constructs are subjective, internal experiences of an individual that do not have a behavioural manifestation (unless their intensity reaches substantially high levels, and that was not the case in this study). Consequently, these constructs were measured based on the subjective reports of the participants. One could question the accuracy of these subjective reports, and particularly the self-reports that could be seen as biased. To mitigate the effects of this issue, a comparative judgment strategy was employed. Participants were asked to compare their problem-solving experiences for the two problems they solved from the same domain in the context of the study rather than some absolute, general scale. This provided a consistent reference point for the participants' judgment and effectively captured the magnitude of the difference in their experiences, which was the key point of interest in this study.

The subjectivity of the judgments used in this study was unavoidable since the phenomena investigated were the inherently subjective and internal experiences of an individual.

The relative intensity of the insight experience was significantly higher in each of the treatment groups compared to the control group, which provided support for Hypothesis 2. Hypothesis 2 states that when the solution to the same problem is found with different levels of tension, the intensity of the insight experience will be greater in the case where a greater amount of tension is released with the solution to the problem. Moreover, there was a highly significant linear correlation ($\rho = 0.73$, $p < 0.01$) between the relative intensity of the insight experience and the relative amount of tension released, further demonstrating the effect of tension on the intensity of the insight experience.

The manipulation of the level of tension also resulted in significantly higher perceived relative difficulty levels and significantly higher relative solution times in the two treatment groups when compared to the control group. Thus, it was demonstrated that without adjusting the problem itself, the manipulation of tension had an effect not only on the perceived difficulty but also the solution times (the more 'objective' measure of difficulty). This finding supports the proposed theoretical framework, which argues that the difficulty of a problem is a function of the amount of tension in the life space and the degree of restructuring required to solve that problem.

The analysis of the tension curves showed that tension during insight problem solving in Experiment 3 had a tendency to increase throughout the process at a decreasing rate. There was no significant decrease in the tension curves observed before the solution point. The largest tension increase took place in the first third of the curves and the least increase in tension took place at the last third of the curves. The diminishing rate of tension increase could be explained by tension adjustment mechanisms that may have been employed by individuals to cope with uncomfortable levels of tension. These mechanisms might involve adjusting the valences of the various sub-regions, leading to a decrease in their force-fields and consequently reducing the rate of the tension increase. This would allow for a reduction in the rate of tension increase, but not for a reduction in the level of tension. At the time such mechanisms are employed, a certain level of tension already exists in the region and needs to be released either through reaching the goal or a substitute activity.

In the two treatment groups, the tension increase did not diminish as much towards the end of the problem solving process, which could be explained by the reduced effectiveness of the tension coping mechanisms. This might be because the adjustment of the valences in the region might be more difficult under stronger force-fields, or because the rate of tension reduction was overall slower that the rate of tension increase.

3.5. Experiment 4

This experiment was motivated by an observation during Experiment 3. From the results of Experiment 1, the benchmark problem of Experiment 3 was expected to be perceived to be more difficult than the experimental problem. The solution time data from the control group of Experiment 3 supported this prediction. In the control group of Experiment 3, the benchmark problem took longer to solve than the experimental problem (Wilcoxon signed ranks test $Z = -1.84$, $N = 40$, $p < 0.05$ 1-tailed). However, in their verbal descriptions, the participants of Experiment 3 often attributed the difficulty of the benchmark problem to the fact that it was the first problem they solved from this domain, and that solving this problem first made the next problem, the experimental problem, easier than it would have been otherwise. In fact, 23 out of 40 participants in the control group of Experiment 3 commented on this. It is worth noting, that this information was volunteered by the participants. Moreover, the participants indicated that solving the benchmark problem first reduced their tension level for the experimental problem.

From the field theory perspective, when a participant is confronted with their first problem, their search space region (the region that separates the initial state from the goal) is far less structured than when they approach their second problem. Before the first problem, the participants do not really know what possible moves could be used in this class of problems or what “tricks” could exist there. When a participant starts working on their first problem, they would be more likely to reach an impasse or a state of conflict sooner since they have to ‘invent’ different possibilities from scratch. However, when they start working on their second problem, they already have a set of different transformations in mind that they tried on their first problem and could be applied to the second problem. As a result, a state of impasse would normally be reached later (if at all) in the process for the second problem than for the first because there are more transformations readily available for the second problem than there were for the first problem. Consequently, the individual would try all these available transformations before running out of ideas and reaching a state of conflict.

Based on this argument, one would expect a faster increase in the tension level in the beginning of the first problem than in the beginning of the second problem.

There could also be a difference in the valence of the goal sub-region between the first and the second problems. The first problem might have a higher valence because an individual has never solved problems of this sort before, and solving the first problem could be associated with proving to themselves that they are capable of solving problems from this domain. After a successful solution of the first problem, the second problem from the same domain is not valued as much, because the participant has already proven that they can solve "these kinds" of problems. The second problem becomes "another one of those," thereby reducing the valence of the goal. Consequently, the valence of the failure region could be higher for the first problem than for the second because of the greater uncertainty associated with the first problem. The changes in the valences of the sub-regions between the first problem and the second problem would only take place if the individual does, in fact, group these two problems as described above. Alternatively, the two problems might be seen as two 'different' problems where solving the first is not related to solving the second one. In this situation, a change in the valences is unlikely, although the structural differences in the region between the first problem and the second problem will still apply.

To summarize, both the structural differences of the search region, and possible differences in the valences of the goal and failure sub-regions, potentially influenced the difference in the strength of the field-forces and differences in the levels of tension between the first and second problems.

From this, one can infer that the higher levels of tension observed for the benchmark problem in the control group of Experiment 3 must have been at least partially due to the fact that it was solved first. If the experimental problem was solved first, then one would expect the difference between the benchmark problem and the experimental problem to diminish.

The order effect of the problems was investigated further. An additional study was conducted to examine how the sequence in which these two problems were solved affected the level of tension, perceived difficulty, and intensity of the insight experience. The study employed a within participant assessment.

3.5.1. Method

Participants. 38 undergraduate students from the University of Waterloo participated in this study in exchange for a partial course credit. Three of the 38 participants could not solve one of the two problems during the experimental session and were excluded from the analysis.

Materials. Same problems as in Experiment 3 were used as stimuli in this study. The problems were represented with brown coffee sticks on the table in front of the participants.

- Benchmark problem: IX = III + I, solution IV = III + I
- Experimental problem: IV = III – I, solution IV – III = I

Procedure. The study was conducted in a laboratory setting with one participant at a time. Each experimental session lasted 1 hour and was videotaped. The procedure employed in this study was the same as for the control group in Experiment 3 with the only difference being that participants first were asked to solve the experimental problem and then the benchmark problem. Only the order of the problems was changed. The measures were the same as in Experiment 3.

3.5.2. Results

Experiment 4 was designed to examine the order effect of the two problems on the levels of tension, intensity of insight and the perception of difficulty. Data from the control group of Experiment 3 was used to make the comparison. In the following discussion, the control group from Experiment 3 is referred to as the “direct sequence group,” and the group in Experiment 4 is referred to as the “reverse sequence group.”

The following experimental predictions were tested. Regardless of which problem was solved first;

- Higher levels of tension will be generated and released solving the first problem;
- Tension on the first problem will increase faster in the beginning of the solution process than on the second problem;
- More intense experience of insight will be associated with the first problem;
- The first problem will be perceived to be more difficult;
- The first problem will take longer to solve.

In the direct sequence group the participants first solved the benchmark problem followed by the experimental problem. In the reverse sequence group, the participants solved the experimental problem first and then the benchmark problem. The measures of the experimental problem of the reversed group were compared to the measures of the benchmark problem of the direct group (i.e. when each problem was solved first in the sequence).

Since the four measures used in this study did not satisfy the conditions of normality and homogeneity of variance, the medians and quartiles are reported in Table 17.

Table 17: Solution times (second), levels of tension, intensity of insight, and perceived difficulty

Problem	Statistic	Measure							
		Solution time		Tension level		Intensity of insight		Perceived difficulty	
		Direct group	Reversed group	Direct group	Reversed group	Direct group	Reversed group	Direct group	Reversed group
1st problem	Low *Q*	41	33	3.9	2.7	4.5	4.2	2	2
	Mdn	108	86	6.6	5.9	7	9	6	5
	Upper *Q*	289	234	8.6	8.7	9.5	10	7	7
2nd problem	Low *Q*	23	19	1.9	2	3	3.4	2	3
	Mdn	61	81	3.5	4.3	4.7	6	3.8	5
	Upper *Q*	168	216	8.1	8.8	9	8.6	6	7

First, the possible differences between the groups were assessed. Single *U* tests were computed to determine the significance of the differences between the groups on all four measures for the first and second problems. The results showed that there was no significant difference on any of the measures between the groups (all *p*s > 0.3, except for the perceived difficulty of the second problem where $p = 0.11$).

Since there were no significant differences between the groups, the data for the first problem from the direct and the reverse sequence groups was combined for all the measures. Similrly, the data for the second problem was also combined for the two groups. The within-participant differences between the first and the second problems were analysed for the combined data set.

Single Wilcoxon signed ranks tests were computed to assess the difference in participants' experience between their first and second problems. The results are reported in Table 18.

Table 18: Results of Wilcoxon signed ranks tests for the difference between the first and the second problem

Parameter	Measure			
	Solution time	Tesion released after the solution	Intensity of the insight experience	Perceived difficulty
Direction	1st > 2nd	1st > 2nd	1st > 2nd	1st > 2nd
Z	-2.03	-2.17	-2.02	-1.16
N	75	75	75	75
p (1-tailed)	< 0.05	< 0.05	< 0.05	0.13

The test results indicated that the first problem indeed took considerably longer to solve, was associated with release of more tension, and resulted in more intense experience of insight than the second problem. However, the perceived difficulty of the first problem was not significantly higher than that of the second.

To examine the differences in the rate of tension increase between the two problems the tension curves were analysed in more detail. The increases in tension level that took place in the first, second, and third intervals of the curves before the solution point were measured in centimetres for each problem. The medians and quartiles are reported in Table 19.

Table 19: Tension curve increases across three intervals before the solution

Problem	Statistic	Tension curve interval					
		Beginning		Middle		End	
		Direct group	Reversed group	Direct group	Reversed group	Direct group	Reversed group
1st problem	Low *Q*	0.7	1.2	0.5	0.7	0	0
	Mdn	1.6	2.4	1.4	1.6	0.6	0.3
	Upper *Q*	3.5	3	2.3	1.9	1.6	1
2nd problem	Low *Q*	0.2	0.3	0	0.1	0	0
	Mdn	0.8	1	0.5	0.9	0	0.5
	Upper *Q*	2.8	2.3	1.6	2.7	1	1.3

The computed U tests showed that there were no significant differences between the groups on any of the intervals (all ps > 0.3), and the data for the first and second problem from the two groups was combined for further analysis.

The significance of the difference in tension curve increases between the first and the second problems was assessed with single Wilcoxon signed ranks tests across the three intervals. There was a significantly higher increase in tension curves in the beginning interval of the first problem than on the second ($Z = -2.25$, $N = 75$, $p = 0.01$, 1-tailed). There was a higher increase in tension during the middle and end intervals of the first problem than on the second, but they did not reach a significant level ($Z = -1.49$, $N = 75$, $p = 0.14$, 2-tailed for the middle interval; and $Z = -1.59$, $N = 75$, $p = 0.11$, 2-tailed for the end interval).

3.5.3. Discussion

This study examined the order effect of the problems on the participants' experience. It was initially suggested by the participants of the previous study, and later explained theoretically, that there is a possible structural and dynamic difference in the life space between the first problem and the second problem.

The novelty and unfamiliarity of the first problem from a domain that one has to solve is associated with a more unstructured search region in the beginning of the problem solving process. This, in turn, leads to an earlier encounter with barriers and, potentially, reaching a state of conflict resulting in a subsequent tension increase. Thus, it was argued that tension would increase earlier on the first problem than on the second – an assertion that was supported by the analysis of the tension curves. There was a significantly higher increase in tension during the beginning interval of the curves for the first problem compared to the second.

The valences of both the goal sub-region and the failure sub-region were potentially higher for the first problem. The valence of the goal sub-region might have been higher on the first problem because of the problem's novelty, and it potentially decreased for the second problem due to familiarity. The higher goal sub-region valence increased the goal-force in the field. The valence of the failure sub-region might have been higher because the problem solver was unsure about whether or not they will be able to solve the problem. The higher negative valence of the failure sub-region potentially also created a stronger force pushing away from that region and contributing to the magnitude of the resultant force in the field.

There was no direct measure of the valences and forces in the life space of the participants; the participants only provided their tension curves and a relative comparison of the intensity of insight and perceived difficulty for the two problems. The above explanation remains on a theoretical level as there is no direct evidence for the adjustment in the valences and forces. The results of this study, however, showed that there was more tension released after solving the first problem than the second, and that this was associated with a more intense experience of insight.

The difference in the perceived difficulty of the two problems did not reach a significant level although the difference in the solution times was significant. This finding was unexpected and could potentially be explained by the combination of the data from the direct and the reverse groups on the difficulty measure. More specifically, the perceived difficulty for the second problem was the only measure that was, although not significantly different between the groups, however only at 0.11 level (compared to other items which were not significant at levels of 0.3 or higher). Perhaps, there were differences in the perceived difficulty judgments in the two groups before the combination which potentially contributed to the variability.

Overall, this study showed that the problem that is solved first has a higher level of tension released after its solution and a more intense experience of insight than the second problem. This finding makes the effects observed in Experiment 3 more powerful, since the manipulation of tension was applied to the second problem of the pair.

4. General Discussion

Insight in problem solving is a fascinating phenomenon of human experience. It was defined by Gestalt psychologists as a sudden restructuring. In the context of this work the phenomenon of insight in human problem solving was examined from a novel perspective - as an experience that has intensity and is affected by the dynamics of the solution process. Insight in problem solving is characterized by an abrupt switch from the state of relatively high tension before the solution to a rapid release of tension after the solution resulting in a relief and a positive feeling.

4.1. Intensity of the insight experience

The results of the four experiments demonstrated that the intensity of the insight experience was affected by the two factors: the degree of restructuring required to solve the problem and the amount of tension released with the solution. The results of Experiment 1 showed that the participants anticipated a higher intensity of the insight experience for the problem that they perceived to be more difficult. The results of Experiment 2 supported these findings in an actual problem solving situation. The problem that involved a greater degree of restructuring was judged as more insightful.

The results of Experiments 3 demonstrated that the increased level of tension led to an increase in the intensity of the insight experience. The manipulation of tension affected not only the intensity of the insight experience but also the perceived difficulty of the problem, and the relative solution time.

Study 4 showed that the problem that was solved first was associated with higher levels of tension and higher intensity of the insight experience due to the order in which the problems were solved. This finding makes the manipulation of tension in Experiment 3 even more significant, since it was applied to the second problem in the pair and yet produced the predicted effects. That is, the order effect was in the opposite direction from the experimental manipulation.

Only three out of 180 participants (1.7%) in Experiments 2, 3 and 4 reported the same intensity of insight for the two problems they solved. The differences in the intensity of the insight experience were not reported randomly. The participants differentiated the intensity of their insight experience systematically depending on the level of tension associated with it and the degree of restructuring in the problem. This shows that the experience of the intensity of insight is psychologically meaningful, and people are able to systematically differentiate it.

4.2. Restructuring in insight problem solving

Restructuring involved in the solution of insight problems is one of the contributors to the intensity of the solution experience. Solution to any problem (insight and non-insight alike) requires certain degree of change in the problem representation. It was proposed that solutions to insight problems have two distinct characteristics with respect to the required restructuring of the problem representation: First, the major amount of change necessary for the solution is achieved in a single psychological step (even if the solution involves several steps, there is one "crucial" psychological step that restructures the representation). Second, the amount of restructuring achieved over one move (the solution move) is significantly greater than in other situations (non-insight problems). Beyond this difference, the insight problems themselves could also vary with respect to the amount of restructuring involved in their solutions.

A measure of restructuring was developed for the domain of matchstick problems to allow for a more continuous measure of the amount of change involved in solutions to different problems from this domain. The preliminary results suggested that the measure of restructuring is a promising approach to measuring the amount of change and that it could be further adjusted to improve its accuracy. The general approach to analyzing the stimuli and counting the amount of change potentially could be applied to other domains.

4.3. Tension in insight problem solving

The major attention in this work was devoted to the understanding and the examining of the psychological tension during the insight problem solving and its effect on the intensity of the insight experience. It was argued that solution of an insight problem is accompanied by a rapid release of tension that was accumulated in one's life space during problem solving process.

Interesting patterns of tension dynamics were uncovered with the analysis of the tension curves reported by the participants. Overall, 360 tension curves were collected in Experiments 2, 3, and 4, two curves from each of the 180 participants. The vast majority of the tension curves (83.3%) represented an overall increasing trend from the beginning to the solution point; 11.7% of the curves exhibited a decreasing pattern from the start to the solution; and 5% of the tension curves showed a relatively stable pattern with no significant changes form the beginning to the solution point. In all of the graphs the solution to the problem followed by a rapid decrease in tension level that was represented by the drop of the tension curve after the solution. The 'rapid' decrease in tension after the solution was confirmed by the time estimates provided by the participants in Experiment 3. The feeling of 'relief' associated with the solution was also reported by the participants in the Preliminary Investigation. These observations overall supported the proposed field theoretical explanation of the dynamics of tension during the insight problem solving.

The patterns of tension change before the solution showed that the increase in the level of tension tended to slow down towards the second half of the solution process when tension was not manipulated. However, when the participants were put under time pressure or offered a reward, the increase in tension was at a steadier rate throughout the whole solution process. These differences could indicate that tension reduction strategies that had been used under the 'ordinary' conditions were not as effective in the conditions of elevated tension. This might have been the case because the tension was increasing at a higher rate than the tension reduction could occur, or maybe the reduction was simply not possible because of the level of tension. In the Preliminary Investigation, 30% of the participants reported some kind of tension reduction actions during their solution of the problem. However, there is no explicit evidence to support the proposed explanations, and they should be treated as rather exploratory.

The results generally supported the proposed dynamics of tension in insight problem solving. The one aspect of tension in insight problem solving that was not examined in this work was the difference in the mechanism of tension release between insight problems and non-insight problems. It was argued in the theoretical discussion that the proportion of tension in the problem solving region that is released with the solution move in insight problems is significantly greater than in non-insight problems. An examination of this issue would involve a comparison of the dynamics of tension during a solution of an insight problem and a non-insight problem which was not the purpose of this study. The main proposition examined in this work was that greater levels of tension release during and immediately after the solution moment lead to higher intensity of the insight experience.

4.3.1. Effect of tension on performance

One of the interesting findings of this study was that the increased tension not only affected the perception of problem's difficulty (i.e. the problem seemed more difficult under higher levels of tension) but also manifested itself in an objective measure of performance – the solution time. When the experimental problem was solved under increased levels of tension it took more time relative to the benchmark problem than when the tension was not manipulated. The level of tension was influenced in two ways. In one case, a time constraint was imposed. In the other case, the participants were offered a reward for solving the problem.

The detrimental effect of stress and pressure on performance have been noted and investigated in psychology. Two general theoretical approaches have been developed to explain the effect of pressure on performance. The emphasis in both of these theoretical directions was on attention. One line of theories ('*self-focused*') explains the detrimental effect of pressure on performance of the skilled tasks by the increased attention to the execution of the skill and the step-by-step execution process, which becomes detrimental since the automation of the execution of a high-level skill is disrupted (e.g. Kimble & Perlmuter, 1970; Lewis & Linder, 1997). The *self-focused* line of explanations is more pertinent to practiced activities and can not be readily applied to insight problem solving since there is no automation involved in insight problem solving.

The other line of theoretical explanations (‘*distraction’*) proposed that pressure-initiated anxiety occupies the working memory and reduces its available capacity for processing the task, thus, diverts attention from the task (e.g. Baumeister & Showers, 1986; Beilock et al., 2004). The *distraction* line of explanation argues that pressure “destructs” one’s attention from performing the task that requires substantial working memory capacity. According to the *distraction* theories, pressure affects performance on tasks with high working memory demand but not on tasks with low working memory demand (e.g. Beilock & Carr, 2005).

It is not exactly clear whether the tasks of Experiment 3 had high working memory demands. In Experiment 3, the participants were working on matchstick arithmetic problems consisting of three number-elements ranging in value from 1 to 9 and involving one operation. Moreover, the problems were constructed with coffee sticks in front of the participants to allow for physical manipulation of the problem elements. It would be reasonable to expect that if one’s working memory was overloaded, then one would physically move the sticks to reduce the load on the working memory. In fact, the problems were set up with physical sticks to allow for observing the sequences of participants’ moves. However, physical manipulation of the sticks was rare in all three groups in Experiment 3; many participants solved the problems in their mind without ever moving the sticks. Since the participants did not try to unload their working memory by physically manipulating the stimuli, then, most likely, matchstick equations combined with the tension manipulations used in Experiment 3 did not place too high of a demand on the working memory. The distraction theories then would not be able to explain the solution time differences observed in Experiment 3 since there was no evidence of working memory overload.

Although distraction and self-focused theoretical explanations might give an account for deteriorated performance under high tension in some problem solving situations, it is not clear how they apply to the matchstick equations of Experiment 3.

Field theory provides a possible explanation as to why and how increased level of tension might have affected performance in a problem solving situation. According to field theory, when the strength of forces acting in the life space increases, it leads to an increase in the level of tension in one's life spaces. Field theory suggests that an increasing level of tension in one's life space might affect the structure of the life space, and, thus, the performance. The extent of the effect of tension on performance depends on the structural properties of the life space, that is, the structure of the task itself, the strength of the boundaries within the region of activity and the level of communication between the sub-regions.

The region that is associated with an increased level of tension acquires a negative valence, which manifests itself in a force away from that region. Thus, increasing level of tension creates pressure on the boundaries within the region of activity, and has a tendency to affect the organization within the region. Higher tension creates stronger urge to act to relieve it. Tension has a tendency to spread to other regions if the boundaries cannot withstand the created pressure. When tension spreads, the boundaries between the sub-regions are wiped out eliminating the differentiation in the region. This makes the search for a solution to a problem more difficult. Lewin (1935) points out that "corresponding to an increase in the general state of tension is a rapid change of occupation" (p. 95). A more rapid change of occupation serves as a mechanism of diffusing the tension; however, the occupation itself also becomes "more superficial" (Lewin, 1935, p. 95).

When this happens in a problem solving situation, an individual might try the same move that did not work before again, or they might try a move that is very similar to the one they already tried. In other words, due to the increasing and spreading tension in the problem solving region, the "quality" of attempted moves might be reduced.

As a result, it is possible that the increase in the level of tension made the problem more difficult by affecting the structure of the problem solving region, thus making it more difficult to find the correct transformation. That is, the dynamic properties of the solution process might affect the structural or cognitive representation of the problem.

There is no direct evidence for this explanation from the collected data. The matchstick equation representation is fairly simple, which allowed the participants to attempt different transformations in their minds without physically moving the sticks on the table. Although, some participants did explicitly report applying the same move over and over again, the post hoc reports of thought process and the sequence of moves most likely are not exactly accurate and might be far more rationalized than the actual process. As a result, there is no direct evidence on the moves that individuals tried in their mind and their sequence.

Previous studies have also reported increased solution times on insight problems when the participants were offered a monitory incentive for their performance (Glucksberg, 1962; McGraw & McCullers, 1979). Camerer and Hogarth (1999) explained these results by a possibility that increased effort in insight problem solving could have led to a persistence with one approach, making the problems more difficult.

From the field theory point of view, monitory incentives potentially might have increased the valence of the goal sub-region for the participants, or increased the negative valence of the failure sub-region. Thus, the strength of the force field also increased leading to an increase in the level of tension. From the field theory point of view, persisting with one approach would be more difficult under increased levels of tension in insight problem solving situation. This is primarily because in insight problem solving there is no way to gradually progress towards the goal. The persistence with one approach constantly results in failure, which, in turn, leads to an increasing tension and results in a negative valence of that region. Increased tension, thus, might lead to an even earlier abandoning of the approach, which is contrary to the Camerer and Hogarth (1999) suggestion.

The explanations of the detrimental effect of tension on the performance in insight problem solving suggested by Camerer and Hogarth (1999) and that based on field theory would lead to similar predictions although through different mechanisms. There is no evidence in the collected data that would allow to support or dispute either of these explanations.

The effect of tension on the structure of the life space, no doubt, needs to be investigated further. This could be done with a stimulus that compels physical manipulation of the items during the solution process, which will make the thought process and the sequence of attempts more "observable." Field theoretical propositions raise interesting and fundamental questions with respect to problem solving, performance, and general effect of tension on human activity.

4.3.2. Tension release versus tension generation

It is worth noting that the main premise tested in this work was the relationship between the amount of tension and the intensity of the insight experience. A positive relationship was established and supported by the results of Experiments 3 and 4. If one assumes that higher intensity of insight is "good," then one may get an impression that problems should be solved under higher levels of tension.

It is worth pointing out, however, that the intensity of the insight experience depends on the amount of tension *released* with the solution, which might be different from the level of tension *generated* during the problem solving process. In other words, there could be more tension generated during the problem solving process than was released when the problem was solved. For example, when an individual is misled to look for the solution in the wrong place and then spends a great deal of time in the 'wrong alley,' whereas having the accurate information would have made the solution to the problem much easier. Upon finally reaching the solution, this individual might perceive a discrepancy between the *experienced* difficulty of the problem (level of tension generated looking for a solution) and the '*actual*' or *potential* difficulty of the problem had one had the right information. In this case, not all the tension generated during the problem solving might be released when the solution is found, leaving the problem solving region to be associated with a certain level of residual tension even after the problem was solved. Dissatisfaction with one's own performance on a successfully solved problem (e.g. Gick & Lockhart, 1995) is evidence of more tension generated on the way to the goal than was released by solving the problem. The residue tension creates dissonance and calls for its reduction.

It is possible that there exists a certain threshold for the amount of tension that can be released solving a particular problem, which might be related to the degree of restructuring involved. If the amount of tension generated during problem solving exceeds that threshold level, not all the tension will be released with the solution, thus resulting in dissonance.

The levels of tension achieved in this study did not allow for observing such a scenario involving excessive tension. The participants were asked to solve two relatively easy, although not trivial, matchstick equation puzzles. The average solution time for these puzzles was around 1.5 - 2.5 minutes. The tension manipulations in the Experiment 3 were fairly ‘mild’ although they created noticeable differences between the groups. The experimental situation (i.e. being observed and videotaped) possibly contributed to the level of tension in the participants to some degree as well, but all participants were well aware of the fact that there was no penalty for not solving the problems. As a result, the experimental situations only created, at best, moderate levels of tension in the participants’ life space.

As was stated before, most of the participants reported a complete release of tension after solving their problems, and the remaining participants indicated a considerable tension decrease which did not reach the zero level. Some participants returned their tension curve after the solution to the same level they started from at the beginning of the experiment, indicating that it is their ‘normal’ tension level, which is higher than zero, or it is their tension level while in the experimental situation. Others pointed out that their tension did not go to zero after solving the first problem because they knew that they needed to solve one more problem. There was no indication of the residual tension from the problem in participants’ verbal descriptions.

It was not possible to investigate the effects of the tension generation-release difference in the context of the reported studies. The question of problem's capacity for tension release is an important topic in understanding human experience in problem solving, and it needs to be investigated.

4.4. *Possible alternative explanation: concept of subjective probability*

One of the earlier conceptualizations of the intensity of the insight experience that I entertained involved the concept of subjective probability. The initial explanation was that the intensity of the insight experience depends on the degree of restructuring and the subjective probability of solving the problem at the moment of solution. Lower subjective probability at the moment of solution contributes to the apparent suddenness of the solution, and, thus, a more intense experience of insight. Although, the concept of subjective probability at a first glance seems to provide a reasonable explanation of the intensity of the insight experience, it was abandoned for a number of reasons.

4.4.1. Subjective probability in the context of problem solving

The first difficulty with applying the concept of subjective probability was a definitional ambiguity. In the context of problem solving the subjective probability represents one's own judgment regarding the likelihood of solving a given problem at some point in the future. The estimation of subjective probability of one's future performance in a given domain has to be based on one's previous experience. For a new domain of activity, i.e. solving matchstick equation puzzles, there is no direct previous experience to draw upon for estimation. Thus, it is not clear how one would go about estimating one's subjective probability.

Also, problem solving is an activity that is associated with certain duration, and has a beginning and an end. How confident one feels regarding the likelihood of solving the problem (in the future) most likely depends on the timeframe of reference, i.e. how much time is available for solving the problem. As a result, in the context of problem solving, the subjective probability has to be defined with respect to certain duration, because this probability might change with the timeframe. For example, consider the concept of subjective probability for these four intervals:

- solving the problem in the immediate time interval, i.e. next move will solve the problem;
- solving the problem in the next two minutes;
- solving the problem within the next twenty minutes;
- solving the problem eventually, given all the time one needs (i.e. take the problem home).

The subjective probabilities associated with the above four timeframes most likely will be different, and will, perhaps, increase from the first to the last. This also implies that one might obtain different judgments of the likelihood of solving a particular problem depending on the timeframe of reference. Thus, the concept of subjective probability needs to be defined with respect to some duration.

When applying the concept of subjective probability to the intensity of the insight experience, the issue of the timeframe of reference remains. The seemingly appropriate timeframe for this context, perhaps, might be the immediate likelihood of solving the problem at the moment of solution. However, the subjective probability might also be defined relative to the context of the study. That is, the subjective probability can be defined with respect to the duration of the study in which a participant needs to solve the problem.

4.4.2. Dynamics of subjective probability during problem solving

The second difficulty with applying the concept of subjective probability is related to its dynamics during the solution process. It is not exactly clear how the subjective probability changes, if at all, during the problem solving process.

On one hand, a problem (e.g. matchstick arithmetic equation) has a limited search space, and with every additional failed attempt there are fewer things left to try. Therefore, the probability of solving the problem should increase with every additional failed attempt as there are fewer things left to choose from. This pattern of change in probability portrays more of an "objective" rather than subjective representation of events in problem solving. However, it is also a plausible direction for change in the 'subjectively' predicted likelihood.

On the other hand, with additional failing attempts and time going by, it might be reasonable to expect a decrease in one's subjective probability of solving the problem which eventually might reach a fairly low level. If this is the pattern of dynamics of subjective probability in problem solving, then there is an issue of what sustains the problem solving activity given a low perceived likelihood of succeeding. In Experiment 2 the participants had to solve a division problem which is relatively more difficult than other problems used. Although, the solution rate for that particular problem was much lower than for the other problems, most of the participants who did not solve the division problem in the context of the study were stopped by the experimenter rather than quitting voluntarily.

Furthermore, every time a person has an idea of a possible move to try, the immediate subjective probability of that move before it was applied, probably, is relatively high, because otherwise why would one try a move if one thinks that the move is going to fail. Later, after the move had failed, the subjective probability of that particular move that failed, no doubt, decreases. However, it is not clear how and why the previous failures affect the subjective probability of the future moves. Further, it is not clear how the person's awareness of the existence of the correct solution to the problem effects the subjective probability estimation.

4.4.3. Psychological representation of the experience

The concept of subjective probability was not adopted also because it does not capture and represent the richness of the psychological experience during the problem solving process. The notion of being uncertain whether or not one can solve the problem during problem solving is, no doubt, sensible. However, by itself it does not capture the essence of the experience of being frustrated and getting stuck on a problem. Just the fact that one's subjective probability of solving the problem is low does not explain the feeling of stress and frustration before the solution and a relief after the solution so frequently reported by the participants in the Preliminary Investigation. In fact, while only three participants of that study commented on the feeling of uncertainty in their ability to solve the problem (which can be interpreted as a low subjective probability), all of the participants commented on the experience closely related to tension. The concept of subjective probability is a high-level construct, and it is not clear if a person actually experiences problem solving process in terms of changing likelihoods of finding the solution. Finally, the concept of subjective probability does not provide a measurement advantage over the concept of tension. The measurement of both concepts has to rely on the subjective judgments of the participants.

Ormerod et al. (2002) rightly pointed out that the state of impasse is more of a state of conflict under high tension; however the concept of subjective probability alone would not have captured this fundamental characteristic of the problem solving experience.

As a result, the direction suggested by Ormerod et al. (2002) was followed in this thesis to analyse insight problem solving as the process that is affected by the dynamics in the state of tension of a person.

4.5. Application of Field Theory to the study of the intensity of the insight experience

Field theory (Lewin, 1935, 1936, 1938) is a comprehensive body of fundamental psychological constructs that represent a method of analyzing, understanding, and predicting human behaviour. This thesis has shown yet another domain of human experience that could be understood better by applying the field theoretical analysis.

In this work, field theory of Kurt Lewin was applied to analyze the insight problem solving and explain the intensity of the insight experience. It allowed for examining the dynamics of insight problem solving process and track the potential changes in the state of tension of an individual. The current cognitive theories of insight could not have provided such an examination. The field theoretical explanation of dynamics of tension was supported by the results of Experiment 2, 3 and 4. Moreover, the reported phenomenological experience of participants in the preliminary investigation was in-line with what the field theoretical explanation would predict.

Application of field theory proved to be valuable in many different areas that Lewin himself was involved. “The breadth of the fields and interests to which Lewin’s thought contributes testifies both to the fundamental nature of many of his ideas and to his breadth as a social scientist…Kurt Lewin is a unique standing model for what contemporary social scientists can be” (Gold, 1999, p. 4 - 5).

4.6. Limitations

The data from Experiment 1 contained some degree of noise. The participants in that study received their experimental booklets by e-mail and completed the activity outside of the laboratory. Although numerous instructions were provided in the booklets, it is possible that the activity might have been completed in a careless manner by some. Despite of these shortcomings it was necessary for testing Hypothesis 1 to remove any possible sources of tension that might be associated with solving a problem and being in an experimental situation.

There were no secondary measures available to validate both measure of tension and measure of the intensity of the insight experience. All of the experimental sessions were videotaped with the hope of observing overt behaviour manifestations of both of these experiences. For example, the increased level of tension might manifest itself in the restless movement (Lewin, 1935, 1938) and the experience of insight might manifest itself in a facial expression. However, during the analysis of the video recordings there were no observable behaviour manifestations found. This could be explained by the very likely possibility that the intensity of the created situations did not reach the level when the experienced tension and insight manifest themselves in the overt behaviour. Also, during the experimental session, the participants were observed and videotaped which might have increased their self-awareness and inhibited manifestation of their experiences.

The measurement of tension and the intensity of insight experience relied on subjective reports after the problems were solved in Experiments 2, 3 and 4. Subjectivity of the measurements could not have been avoided in this study as the nature of the studied phenomena is subjective experiences of an individual and no behavioural manifestations of these experiences were observed. A care was taken to obtain the measurements as close to the solution of the problems as possible to ensure a greater accuracy, however, it did not eliminate the after-the-fact reporting. A possible bias in the subjective reports could have been due to recency of the experience, that is the last problem that was solved being more vivid in one's memory might have resulted in higher ratings of the intensity of insight or tension. However, it was not the case in the reported studies as the difference between the control group and the treatment groups indicated.

The study that utilized a between-participant design, Experiment 3, was designed to obtain a comparative judgment. The participants were taken through two similar experiences (i.e. solving two problems from the same domain) and were asked to compare their experience in those two situations rather than judge their experience on some absolute, general scale. This provided a consistent reference point for participants' judgment and allowed to capture the magnitude of the difference in their experience. The interest of the study was to investigate the difference in the participants' experience.

Besides the measurement advantage that the comparison provided, solving two problems in the context of the study introduced additional variability due to the possible learning effects.

This study used field theoretical concepts to explain the phenomenon of insight. These concepts (e.g. forces and valences) are not directly observable, but only their effects manifested in behaviour might be observed. The fact that psychological forces are not directly observable does not undermine their theoretical soundness - physical forces are not observable either, only their effects are. However, psychological forces and valences can not yet be directly measured to calculate the resultant force and predict the behaviour.

Many theoretical explanations derived from the field analysis were not tested, such as the effect of tension on the structure of the life space, tension reduction mechanisms used during problem solving and how they operate, the potential discrepancy in the levels of tension generated and released during problem solving. These will need to be addressed in future research efforts.

4.7. Future research

There are three potentially promising directions for future research that emerged from this study. First, the proposed measure of restructuring in matchstick arithmetic domain could be analyzed further, validated, and, possibly, extended to other domains. Second, it might be valuable to investigate the discrepancy between the tension accumulation during the problem solving and its release after the solution (i.e. more tension generated than tension released at the moment of solution). Third, more research is needed to examine the effect of tension on problem structure and performance in insight problem solving. More specifically, how high tension affects the selection of moves or transformations during the problem solving.

The measure of restructuring developed for the domain of matchstick arithmetic problems was not the main emphasis in this work, and it did not receive substantial attention. This measure was verified only in the context of a preliminary study, and more could be done, undoubtedly, to improve it, obtain further evidence, and possibly extend the general approach to other domains of problems. It seems plausible to apply the general approach of this measure to other problems that involve several interrelated elements, such as geometric shape stick puzzles and coin puzzles. Many of these problems involve more than one move, and the measure might allow for comparing the relative degree of restructuring of the moves within the same problem. The availability of a continuous measure of restructuring for a series of problems has number of advantages to the study of problem solving as was outlined in section 2.1.3.2., and development of such measure need not be neglected.

It was observed in previous studies of insight problem solving that sometimes participants expressed dissatisfaction with their performance or the problem after solving it (usually after a prolonged series of attempts) (e.g. Gick & Lockhart, 1995). Such reaction indicates an existing tension in the life space of an individual that was not released with the solution. It was suggested above, that there may be a threshold for the amount of tension that could be released for a given problem which is related to the degree of restructuring of the problem. If more tension was generated than released after the solution, then the dissonance will occur. This phenomenon could be examined using the fixation effect. Following the analogy of the water jar problems of Luchins (Luchins & Luchins, 1950) a similar set could be constructed from the matchstick arithmetic problems. Participants might be asked to solve a set of problems that all require a division solution and then presented with a problem that requires a much simpler solution and can not be solved with a division operation. This potentially will create a mental set for the last problem which will perhaps be solved with greater difficulty than it would have been otherwise, thus possibly allowing observing the difference between the amount of tension generated before the solution and the amount of tension released after the solution.

Another line of research could be developed investigating the effect of tension on the structure of the life space during insight problem solving and its effect on the solution process. This requires a possibility of observing the solution process as it unfolds. One potential set of stimuli that could be used in such investigation is spatial puzzles such as Tangram or Soma cube that require manipulation of the physical puzzle pieces to construct a certain shape. In these puzzles, it is necessary to move the pieces in order to evaluate the possible solution paths. One could analyze the stimulus and develop a measure of similarity of different moves, and then observe the solution process under different conditions.

Conclusions

This work examined the dynamic aspect of the phenomenon of insight in problem solving and how it affects the intensity of the insight experience. The field theory of Lewin (1935) proved to be quite instrumental in providing an explanation of the psychological experience during the insight problem solving process.

The insight problem solving process was conceptualized as the process that is characterized by an increasing tension in the life space before the solution. The solution of an insight problem is characterized by a greater amount of change in the problem representation than in non-insight problems and is accompanied by a rapid release of tension.

The intensity of the insight experience was conceptualized as a function of the degree of restructuring and the amount of tension released after the solution. The reported studies provided support for this conceptualization showing that when either the degree of restructuring or the amount of tension release were varied independently from each other, it led to higher intensity of insight.

It is also worth noting that both the degree of restructuring and the level of tension are not acting independently from each other. In an actual problem solving situation, the degree of restructuring affects the level of tension, and on the other hand, the level of tension affects the structure of the life space making the restructuring easier or more difficult. Four experiments were conducted to examine the effect of these two factors and their results provided an empirical support for the theoretical explanation.

The present work has made three potential contributions to the field. First, a more continuous measure of restructuring was developed for the domain of matchstick problems. Based on the preliminary testing of this measure, it seems to provide an improved prediction over the predictions that could have been obtained with the existing problem taxonomy (Knoblich et al., 1999). Second, the notion of the intensity of the insight experience was introduced and analyzed as an abrupt release of tension that accompanies restructuring, representing a novel view of the phenomenon of insight. Third, field theory (Lewin, 1936) was applied to analyze insight problem solving experience. This analytical approach allowed for conceptualizing the cohesiveness of both the structural and the dynamic aspects of the insight problem solving experience and their interrelationship within the same framework.

References

Ash, I. K., & Wiley, J. (2004). Ah-Ha! I knew it all long. Hindsight bias evidence of restructuring in problem solving. Paper presented at the *Annual Meeting of the Midwestern Psychological Association*: Chicago, IL. Available from the author's website http://litd.psch.uic.edu/personal/jwiley/Ash_wiley.pdf (accesses on June 05, 2004).

Ash, I.K. & Wiley, J. (2006). The nature of restructuring in insight: an individual-differences approach. *Psychonomic Bulletin & Review, 13 (1), pp. 66 – 73.*

Back, K. W. (1992). This business of topology. *Journal of Social Issues*, *48 (2)*, pp. 51 - 66.

Baumeister, R. F. & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology, 16, pp. 361 – 383.*

Beilock, S. L. & Carr, T. H. (2005). When high-powered people fail. *Psychological Science, 16 (2), pp. 101 – 105.*

Beilock, S. L., Kulp, C. A., Holt, L. E., & Carr, T. H. (2004). More on the fragility of performance: Choking under pressure in mathematical problem solving. *Journal of Experimental Psychology: General, 133 (4), pp. 584 – 600.*

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57 (1), pp. 289 – 300.*

Benjamini, Y. & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics, 29 (4), pp. 1165 – 1188.*

Bowden, E.M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences, 9 (7), pp. 322 – 328.*

Chronicle, E. P., MacGregor, J. N. & Ormerod, T. C. (2004). What makes an insight problem? The roles of heuristics, goal conception, and solution recoding in knowledge-lean problems. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30 (1)*, pp. 14 – 27.

Chronicle, E. P., Ormerod, T. C. & MacGregor, J. N. (2001). When insight just won't come: The failure of visual cues in the nine-dot problem. *The Quarterly Journal of Experimental Psychology, 54A (3),* pp. 903 – 919.

Camerer, C.F. & Hogarth, R.M. (1999). The effects of financial incentives in experiments: a review and capital-labor-production framework. *Journal of Risk and Uncertainty, 19 (1-3); pp. 7 – 42.*

Chrysikou, E.G. & Weisberg, R.W. (2006). Following the wrong footsteps: fixation effects of pictorial examples in a design problem-solving task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(5), pp. 1134 – 1148.*

Davidson, J. E. (1995). The suddenness of insight. In R. J. Sternberg & J. E. Davidson, Eds. *The Nature of Insight*. The MIT Press. pp. 125 - 156.

Davidson, J. E. (2003).Insights about insightful problem solving. In J. E. Davidson & R. J. Sternberg, Eds. *The psychology of problem solving*. Cambidge University Press, pp. 149 - 175.

Dominowski, R. L. & Bourne, L. E., Jr. (1994). History of research on thinking and problem solving. In R. J. Sternberg, Ed. *Thinking and Problem Solving*. San Diego: Academic Press, pp. 1 - 36.

Dominowski, R. L. & Dallob, P. (1995). Insight and problem solving. In R. J. Sternberg & J. E. Davidson, Eds. *The Nature of Insight*. The MIT Press. pp. 33 - 62.

Dunker, K. (1945). On problem solving. *Psychological Monographs. 58, 5, whole No. 270.*

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press

Gick, M. L. & Lockhart, R. S. (1995). Cognitive and affective components of insight. In R. J. Sternberg & J. E. Davidson, Eds. *The Nature of Insight*. The MIT Press. pp. 197 - 228.

Gilhooly, K.J. & Murphy, P. (2005). Differentiating insight from non-insight problems. *Thinking and Reasoning, 11(3), pp. 279 – 302.*

Glucksberg, S. (1962). The influence of strength of drive on functional fixedness and perceptual recognition. *Journal of Experimental Psychology, 63, 36 – 41*.

Gold, M. (1990). Two "Field Theories". In S. A. Wheelan, E. A. Pepitone, and V. Abt, Eds., *Advances in Field Theory*. London: Sage Publications, pp. 67 - 79.

Gold, M. (1992). Metatheory and field theory in social psychology: Relevance or elegance? *Journal of Social Issues*, *48 (2), pp. 67 - 78.*

Gold, M. (1999). Introduction. In M. Gold, Ed., *The Complete Social Scientist*. Washington, DC: American Psychological Association, pp. 3- 5.

Grant, E. R. & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14 (5)*, pp. 462 – 466.

Hall, C. S. & Lindzey, G. (1978). *Theories of Personality*. New York: John Wiley & Sons.

Jones, G. (2003). Testing two cognitive theories of insight. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29 (5)*, pp. 1017 – 1027.

Kaplan, C. A. & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22*, pp. 374 – 419.

Katona, G. (1940). *Organizing and memorizing: Studies in the psychology of learning and teaching*. New York: Columbia University Press.

Kershaw, T. C. & Ohlsson, S. (2004). Multiple causes of difficulty in insight: The case of the nine-dot problem. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30 (1)*, pp. 3 – 13.

Kimble, G. & Perlmuter, L. (1970). The problem of volition. *Psychological Review, 77, pp. 361 – 384*.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25 (6),* pp. 1534 – 1555.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory and Cognition, 29 (7)*, pp. 1000 – 1009.

Kohler, W. (1925). *The Mentality of apes*. New York: Liveright.

Lewin, K. (1935). *A Dynamic Theory of Personality: Selected Papers*. New York: McGraw-Hill Book Company.

Lewin K. (1936). *Principles of Topological Psychology*. New York: McGraw Hill Book Company.

Lewin, K. (1938). The conceptual representation and measurement of psychological forces. In *Contributions to Psychological Theory,* Vol. 1, No. 4, Durham, NC: Duke University Press.

Lewin, K. (1951). *Field Theory in Social Science*. D. Cartwright (Ed.). New York: Harper & Brothers Publishers.

Lewis, B. P. & Linder, D. E. (1997). Thinking about choking? Attentional Processes and paradoxical performance. *Personality and Social Psychology Bulletin, 23(9), pp. 937 – 944*.

Luchins, A. S. & Luchins, E. H. (1950). New experimental attempts at preventing mechanization in problem solving. *The Journal of General Psychology, 42*, pp. 279 – 297.

Luchins, A. S. & Luchins, E. H. (1959). *Rigidity of Behaviour*. University of Oregon Books, Oregon.

McGraw, K.O. & McCullers, J. C. (1979). Evidence of a detrimental effect of extrinsic incentives on breaking a mental set. *Journal of Experimental Social Psychology, 15, pp. 285 – 294.*

MacGregor, J. N., Ormerod, T. C. & Chronicle, E. P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27 (1),* pp. 176 – 201.

Mayer, R. E. (1983). *Thinking, Problem Solving, Cognition*. New York: W. H. Freeman and Company

Mayer, R. E. (1995). The search for insight: Grappling with Gestalt Psychology's unanswered questions. In R. J. Sternberg & J. E. Davidson, Eds. *The Nature of Insight*. The MIT Press. pp. 3 - 32.

Metcalfe, J. (1986a). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *12 (2)*, pp. 288 – 294.

Metcalfe, J. (1986b).Premonitions of insight predict impeding error. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *12 (4),* pp. 623 – 634.

Metcalfe, J. & Wiebe, D, (1987). Intuition in insight and non-insight problem solving. *Memory and Cognition, 15 (3)*, pp. 238 – 246.

Ohlsson, S. (1984a). Restructuring revisited I. Summary and critique of the Gestalt theory of problem solving. *Scandinavian Journal of Psychology, 25*, pp. 65 – 78.

Ohlsson, S. (1984b). Restructuring revisited II. An information processing theory of restructuring and insight. *Scandinavian Journal of Psychology, 25*, pp. 117 – 129.

Ohlsson, S. (1992). Information-processing explanation of insight and related phenomena. In M.T. Keane & K. J. Gilhooly (Eds.), *Advances in the Psychology of Thinking* (vol. 1, pp. 1 – 44). London: Harvester Wheatsheaf.

Ormerod, T. C., MacGregor, J. N., & Chronicle, E. P. (2002). Dynamics and constraints in insight problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *28 (4),* pp. 791 – 799.

Scheerer, M. (1963/1967). Problem Solving. In A. J. Riopelle, Ed. *Animal Problem Solving*. Penguin Modern Psychology. Pp. 26 - 42. First published in *Scientific American*, (*1963), v. 208*, pp. 118 - 128.

Schooler, J. W., Fallshore, M., & Fiore, S. M. (1995). Epilogue: Putting insight into perspective. In R. J. Sternberg & J. E. Davidson, Eds. *The Nature of Insight*. The MIT Press. pp. 559 - 588.

Segal, E. (2004). Incubation in insight problem solving. *Creativity Research Journal, 16 (1)*, pp. 141 – 148.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R. J. Sternberg & J. E. Davidson, Eds. *The Nature of Insight*. The MIT Press. pp. 65 - 124.

Smith, S. M. (1995). Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In R. J. Sternberg & J. E. Davidson, Eds. *The Nature of Insight*. The MIT Press. pp. 229 - 252.

Smith, S. M., Ward, T. B., & Schumacher, J. S. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition, 21, pp. 837–845*.

Wallas, G. (1926). The art of thought. New York: Harcourt Brace Jovanovich.

Weisberg, R. W. & Alba, J. W. (1981). An examination of the alleged role of "Fixation" in the solution of several "Insight" problems. *Journal of Experimental Psychology: General*, *110 (2)*, pp. 169 – 192.

Weisberg, R. W. (1992). Metacognition and insight during problem solving: Comment on Metcalfe. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18 (2),* pp. 426 – 431.

Weisberg, R. W. (1995). Prolegomena to theories of insight in problem solving: a taxonomy of problems. In R. J. Sternberg & J. E. Davidson, Eds. *The Nature of Insight*. The MIT Press. pp. 157 - 196.

Wertheimer, M. (1925/1969). Gestalt theory. In W. D. Ellis *A source book of Gestalt Psychology*. London: Routledge & Kegan Paul Ltd.

Wertheimer M. 1959. *Productive Thinking*. Harper & Row, NewYork.

Wieth, M. & Burns, B. D. (2000). Motivation in insight versus incremental problem solving. Paper presented at the Twenty Second Annual Meeting of the Cognitive Science Society, Philadelphia, Pennsylvania, August.

Wieth, M. & Burns, B.D. (2006). Incentives improve performance on both incremental and insight problem solving. *The Quarterly Journal of Experimental Psychology, 59 (8), pp. 1378 – 1394.*

Appendix A: A short summary of Lewin's theory

"The basic idea of the field theoretical approach is that every behaviour is a function of the total life space which includes both the person and the environment" (Lewin, 1938, p. 96). Field theory for studying individual psychology developed by Kurt Lewin offers many advantages for analysing psychological phenomena. In his theory Lewin viewed behaviour as the resultant of the properties and dynamics of one's psychological field "here and now," as a product of totality of the psychological situation. "Every psychological event depends upon the state of the person and at the same time on the environment" (Lewin, 1936, p. 12). The psychological field includes both, the person and the psychological environment of that person. As a result, the basic formula of the field theory is

$$B = F\ (P, E) \qquad \text{(Lewin, 1938)}$$

This expression postulates that behaviour (B) is a function (F) of the person (P) and the environment (E). This Lewin's principle is now widely accepted in psychology (Marrow, 1969). Both, the individual and the environment are treated as interdependent factors. Dynamic aspects of the immediate experience caused by the interrelationships of forces and other dynamic factors present in one's psychological field received special attention in Lewin's theory.

Life space

The only determinants of behaviour at a given time, according to Lewin are the properties of the field at that time. The psychological field of an individual is called life space. "Life space is the total psychological environment which the person experiences subjectively" (Marrow, 1969, p. 35). Psychological life space indicates "the totality of facts which determine the behaviour of an individual at a certain moment" (Lewin, 1936, p. 12). All psychological events occur within the life space. Person's immediate situation is represented as a position in a region of activity in a life space. Region of activity refers to a psychologically meaningful qualitative unit of activity in which a person has a place and in which he/she moves. For example, possible regions of activity could be “making breakfast”, “listening to a lecture”, “making summer vacation plans”, etc. Each region of activity is a part of a more inclusive whole, with life space itself being the most inclusive region. For example, region of “making breakfast” could be a part of a more inclusive region of “getting ready in the morning”. Each region of activity could be at the same time divided into several sub-regions that portray the psychological structure of that region of activity in more detail. For example, a region of activity “making breakfast” could include sub-regions “boiling water for tea”, “toasting bread”, and “feeding the cat”. However, the division of a region into sub-parts is not infinite, life space is said to be only finitely divided into regions that are psychologically meaningful to the individual in question. For example, it might not be psychologically meaningful to represent as a sub-region each minute movement of a person’s hand when pressing a toaster leaver. The number of subparts within a region represents the degree of differentiation of that region. Topological representation of life space and regions within it was used to represent an immediate psychological situation of a person.

Boundaries and barriers

Regions of activities within a life space are said to be separated by boundaries. These boundaries could be physical or psychological in nature. For example, an adult’s prohibition of a particular action acts as a psychological boundary of free movement for a child; a boundary of a soccer field is an example of a physical boundary for the players. Boundaries might create a different degree of resistance for locomotion between regions, with the extreme case of impassable boundaries.

Boundaries that offer a resistance to psychological locomotion are called barriers. A barrier can cause different kinds of resistance that can vary depending on the different kinds of locomotion (e.g. physical, social etc.). The resistance to locomotion offered by a barrier is not necessarily constant for the whole region (the barrier might be present at the entrance into a region from one particular region but absent at the entrance from a different region). Also, resistance can be different for various directions of locomotion. For example, a boundary between region *a* and region *b* can offer a great resistance for locomotion from *a* to *b*, however offer no resistance for locomotion from *b* to *a*. A barrier is impassable when locomotion through it is impossible.

Locomotion and change in life space

According to the field theory, at any moment in time, a person is psychologically located in a certain region of activity in their life space. This position is not stationary though, it constantly changes. Locomotion refers to movement of the person through the life space from one region of activity to another or between sub-regions within the same region of activity. The locomotion is said to be steered by the totality of all the forces present in the field at that moment. Objects in one's environment are also represented as regions in a life space. However, the locomotion within or through these regions is not possible. The locomotion through a life space at the same time changes the life space and its properties, i.e. it changes the totality of possible behaviours (Lewin, 1936). With locomotion from one region to another, the magnitude and direction of forces acting on an individual might change, the state of tension might change, locomotion to other regions adjacent to this one might become possible, etc. More drastic structural changes to the life space could also happen. Some of them are differentiation, integration, and restructuring of a region (Lewin, 1936).

- Differentiation of a region in a life space happens when a region is seen as being composed of a greater number of part regions. The differentiation of the life space, for example could happen as a result of gaining experience.
- Integration of a region in a life space refers to a reduction of the number of part regions that compose it. This could happen as the effect of strong tensions (e.g. certain emotional situations), boundaries between certain regions in such situations are said to be destroyed. Integration can also be a result of intellectual processes.
- Restructuring of the life space refers to the change in the relative position of part regions without changing the number of the part regions. Restructuring is often accompanied by either differentiation or integration.

Tension

The concept of tension is one of the central in Lewin's theory and it is closely related to the concept of need. Lewin (1938) writes: "whenever a psychological need exists, a system in a state of tension exists within the individual" (p. 99). "Lewin held that tensions arise when there is a need or want. It is their striving for discharge that supplies energy for, and is consequently the cause of, all mental activity" (p. 30 -31, Marrow, 1969). Tension refers to a state of a region in a life space of an individual, a state of readiness or a preparation for an action. As soon as tension arises there is also a need to relieve this tension which sustains the goal-directed activity until it is completed. As a need is satisfied or a desired goal is achieved, tension is said to be relieved.

"A person might be considered as a system of dynamically more or less inter-dependent sub-systems... One of the outstanding dynamic characteristics of such a system is its tension." (p.97, Lewin, 1938). Term system in this case refers to "a region considered in regard to its state, especially to its state of tension" (p.218, Lewin, 1936). Various regions of activity have different degrees of communication between each other. Degree of communication between two regions refers to the extent to which change in the state of one region changes the state of the other region. Psychological systems within a person have various degrees of communication; as a result, multiple tense systems might co-exist within the same life space without influencing each other, which makes an "ordered action" possible (Lewin, 1935). For example, one's state of tension in professional region might not affect the level of tension in one's family region if these regions have a low degree of communication, or in Lewin's terms, have a "firm wall" separating them. However, when tension in a region reaches a certain level it might spread to other regions when the "walls" can not withstand it any more. Very high tension in one system might result in "flooding" other systems, in this case the spreading of tension occurs and the boundaries between different regions are wiped out. The possibility of spreading of tension depends on the strength of the walls between regions. It is worth noting that locomotion and communication are two different processes between two regions. Locomotion from one region to another might be very well possible; however these regions might have very low degree of communication.

There is a tendency toward an immediate discharge of tension as soon as it arises which could be achieved through a fulfilment of a need or a wish. If such fulfilment is not immediately possible, the system remains in a state of tension. It was experimentally shown in studies of Ovsiankina (reported in Lewin1935, Marrow, 1969) that unfinished tasks remain as tense systems in the life space. These tense systems might have no effect on other systems; a person might go on with other tasks unaffected by the presence of other tense systems. Nevertheless, when an opportunity to complete the previously interrupted task presents itself, the task is resumed, completion of which leads to release of tension. The interrupted tasks were shown not to be resumed if the tension was already released. The voluntary tendency to resumption of an unfinished before task demonstrates that tension remained in that psychological system. Experimental investigations of Zeigarnik (reported in Lewin, 1935; Marrow, 1969) showed that uncompleted tasks are remembered much better than completed ones, which was explained by persisting tension.

Sometimes even when the fulfilment of the need or wish did not occur, discharge of tension may eventually happen through a completion of substitute activity which happens through discharge of tension in a different system with high degree of communication with this one (experiments by Mahler and Lissner reported in Lewin, 1935). In other cases without fulfilment of a wish or need only reduction of tension is possible at best through substitute activities. If tension was not discharged it would remain in the system and influence other systems in a life space only to the extent of the degree of communication between these systems.

One of the important properties of a system in a state of tension is that it "tries to change itself in such a way that it becomes equal to the state of its surrounding systems" (p. 98, Lewin, 1938). Whenever a tension of the present region is higher than that of a neighbouring region, there exists force in the direction toward the region with less tension. Accordingly, tension has a tendency to spread over the whole person especially if it is high (a greater force is created) which is resisted by the firmness of the walls separating the regions.

The extent of tension in a region depends on the strength of the opposing forces. The stronger the opposing forces the greater the tension in the region. Tension increases with encountering a barrier, because the opposition of the forces becomes stronger, increasing even more upon reaching a state of conflict (discussed below).

Valence

Regions of activities within a life space are said to possess certain valances. Valance refers to the notion of the "attractiveness" of a region, and therefore could be positive as well as negative. Valences that certain regions and objects possess at a given moment depend on needs that one has at that moment. Something that had a highly positive valence an hour ago, might acquire a negative valence now due to satiation of a need. Valance of a region generates force acting on an individual. In the case of a positive valance, the created force attracts to that region. If valence is negative, then the created force repels from that region. Positive and negative valences also might have different strength. Strength of the valence determines the strength of the force it generates.

Forces

Forces present in a life space are psychological, not physical and result from tense systems and valences of the regions within the life space. Psychological force is defined as a cause of change. Psychological forces as regarded by Lewin are vectors that have direction, strength, and point of application. Lewin defined driving forces as those that are associated with positive or negative valences. A positive valence of a region creates force directed toward this region; a negative valence of a region creates force directed away from that region. Barriers that impede locomotion are said to exert restraining force on encounter.

Strength of psychological force as defined by Lewin depends on the strength of the valance of the region and the psychological distance to that region (experimentally supported by studies of Fajans reported in Lewin 1938, Marrow 1969).

There are several forces present in a field at any moment in time and their collective effect on the actual behaviour is denoted by the concept of resultant of forces. Lewin (1938) states that "…*an actual locomotion can be related only to the totality of forces acting on a given region at a given time*; in other words, to a "resultant" of forces." (p. 83). The resultant force is the one that steers the behaviour. The notion of resultant force as the drive of the change is central to a concept of the state of equilibrium.

Equilibrium

Equilibrium is such a state where the resultant force is equal to zero, or put in different words when all forces active in the field cancel each other. All psychological processes are steered in the direction of the state of equilibrium (Lewin, 1935). The transition from a state of rest to a process is due to the disturbance of the equilibrium at some point. As a result, a process starts in the direction of a new state of equilibrium. The tendency toward equilibrium is true for the system as a whole, not for particular components. State of equilibrium does not mean that the system is without tension. System can come to equilibrium in a state of significantly high tension (e.g. reaching an impasse in the problem solving situation, or a physical example of a container filled with gas under pressure) this presupposes a certain firmness of boundaries and actual segregation of the system from its environment. If the system is not segregated enough or can not withstand the pressure of the tensions, diffusion into the neighbouring regions will occur, and the equilibrium disturbed. The need to relieve tension in the system can also be viewed as a tendency toward equilibrium, since satisfaction of a need will relieve the tension and minimize driving force and restraining forces due to that need, thus leading to a state of equilibrium at a lower level of tension in the system.

Conflict

A psychological state of conflict is a special case of a state of equilibrium. Psychological conflict is defined by Lewin as the opposition of equally strong field forces. A state of conflict is also a state of high tension in the system. Lewin (1935) differentiates among three basic cases of conflict: conflict between two positive valences; conflict between two negative valences; and conflict when positive and negative valences are in the same psychological location.

Positive – positive conflict happens when a person finds oneself in between two positive regions exerting equal positive forces in the opposite direction. Such state of conflict, according to Lewin, is labile equilibrium, because any movement in the direction of any of the two regions would lead to continuation in that direction.

Negative-negative conflict would happen when a person finds oneself in between two negative regions exerting equal forces in opposite directions and no other movement (except toward one of these two regions) is possible. Situation of threat of punishment if an undesirable task will not get done is an example of this type of conflict if both the task and the punishment are equally negative to an individual and there is no other way for the individual to get out of the situation except through either performing the task or receiving the punishment. This type of conflict, according to Lewin, is stable equilibrium, because any movement in the direction of either of the negative regions will result in return to the original state.

Positive – negative conflict happens when a region with positive valence can only be reached through a region with equally strong negative valence, or when a region with positive valence is surrounded by an impassable barrier. An example of this type of conflict is when an impasse is reached during problem solving, which means that the solution or the goal region of the problem is surrounded for an individual with an impassable barrier that exerts a restraining force equal to the driving force. This type of conflict also is a case of stable equilibrium because any movement away from the barrier would produce resultant force in the direction toward the positive valence, thus, back to the original spot, against the barrier. This explains prolonged attempts at reaching a goal through the barrier although the initial effort was not successful. Before an encounter, a barrier does not have a negative valence as such. However, the barrier itself acquires a negative valence upon an encounter with it that leads to a failure to reach the goal. The strength of this negative valence increases with repeated encounters and continuation of attempts. Lewin (1938) reports experimental studies conducted by Fajans in support of this assertion (p. 127 - 130). The negative valence of the barrier creates an additional force in the field in the direction away from the barrier, therefore, disturbing equilibrium. It leads the person to go out of this region, stop the attempts and do something else. After some time the negative valence of the barrier weakens and attempts might resume.

The described dynamics of these cases of conflict would hold only if the field does not change i.e. all other factors in the life space remain unchanged. As soon as field changes, e.g. other forces arise due to changed valences or constructed regions, resultant force might change and the system will no longer be in a state of equilibrium and conflict. The process will strive to a new state of equilibrium.

Detour

Detour is a round-about way to reach a goal. It occurs when an individual has to find a way around an encountered barrier to be able to reach a goal. A barrier might be physical or psychological. If a barrier is encountered on the way to the goal, then the initial path to the positive valence was perceived to lie in the direction of the barrier, which also means that at that moment any movement away from the barrier also corresponds to the movement away from the positive valence. The difficulty of the detour is in the fact that it requires making a movement in the direction that at that moment seems as being away from the goal. When the detour is found it happens "by reason of a restructuring of the field. There occurs a perception of the total situation of such a kind that the path to the goal becomes a unitary whole. The initial part of route, which objectively is still a moment away from the goal, thereby loses psychologically that character and becomes the first phase of a general movement toward the goal" (p.83, Lewin, 1935). The restructuring of the field can only happen if the psychological field is wide enough to include the whole path. After a while a barrier itself acquires a negative valence, thus pushing away from itself. This might lead to a complete abandoning of the situation and going into a different, unrelated region of activity. Or, realisation on the part of the person that locomotion through the barrier does not correspond to the movement toward the goal might lead to a restructuring of the situation in such way that the path to the goal does not go through the original barrier. The result of this restructuring might lead to an encounter of a new barrier or to the attainment of the goal.

Increased valence, which also corresponds to greater initial tension, makes the solution of detour problem more difficult because an individual has to overcome a greater force toward the goal in order to start moving away from it. Insight problems can be conceptualised as detour problems, with psychological barriers that require restructuring of the field for their solution.

Appendix B: List of Roman numerals from 1 to 12 and mathematical operations

Roman numerals from 1 to 12

I = 1

II = 2

III = 3

IV = 4

V = 5

VI = 6

VII = 7

VIII = 8

IX = 9

X = 10

XI = 11

XII = 12

Mathematical operations

– + x / =

Appendix C: Descriptions of the solution moment by participants in Preliminary Investigation

Below are the excerpts from the responses of the Preliminary Investigation study participants to the question "Could you please describe your experience at the very moment you found the solution to the problem?"

- It's like "BOOM!", it's like an Epiphany, like you have a thought "O, sh.t! This is the solution!" It was a genius moment!
- As soon as I realized what I had to do, everything seemed to flow better. It was a feeling of Epiphany, "O WOW!" feeling.
- I got excited when I solved it, it's like euphoria, a triumphant moment, a moment of pride.
- It was a relief and happiness. Elation!
- Happy, relief, smart, elated, stress free.
- Excitement, sense of accomplishment, feeling of success, I got the initial "YEY! I solved the problem!" when I realized the solution, the excitement peaked.
- I was excited to solve the problem, once I saw it, it was ok, and I wasn't excited anymore, I had excitement for a short while. You are happy when you solve it for a second, but after a second it goes away and you are not excited anymore. Right when you realize the solution you get happy and excited, but then it goes back down.
- It gave me gratification that I solved something. There was a rush which you get when you solve it and then it disappears. That rush that you feel is proportional to the amount of time you spent solving the problem. So, if I've spent an entire afternoon solving something and then I got the solution, then, I think, I would feel more of a rush and it would last longer.
- Unbelievable-surprise, relief, pretty surprised, happy.
- "I got it!" I felt pride that I could think outside the box, feel good about myself, clever.
- Sense of relief; really relieved, a jolt of excitement and confidence, a lot of stress released.
- WOW! That easy?! Surprise, "YEY! I did it!" excited

- “Oh, yeah! Yes! I got it!” relieved, happy, when I picked that stick it was the most intense feeling of happiness and then it levelled off.
- It relaxes you right away, feel elated, you are excited, it’s a good feeling, it’s a relief.
- Happy, relief, satisfaction, it’s like a spark of “YEY! I got it” this spark depends on frustration, when I get something that took me a whole day to figure out, the spark would be much greater than for this problem.
- It feels good, I got it! Feeling of inner strength, lifted feeling. Feeling of success. Jolt of excitement, kind of an “aha!”
- Oh, yeah!!! That’s it!!! At that moment I was relieved and happy.
- I could feel the pressure drop, I could feel my shoulders going from being tense and up to down and relaxed. Relief was a big feeling. There was a definitely rush of excitement, “I got it!!!” It’s definitely exciting. It just clicked!
- A Eureka moment. It was a sense of accomplishment. Feeling of pride.
- A relief, I successfully solved it, self pride, huger reward; euphoria; there was more relief, “YES!” feeling, excitement.

Appendix D: Example of a form sent to the participants in Experiment 1 (reduced size)

MSCI 211 Bonus Research participation Activity

Title of Project: Intensity of Insight in Problem Solving: Structural and Dynamic Properties

Principal Investigator: Professor Frank Safayeni
University of Waterloo, Department of Management Sciences
(519) 884-4567 Ext. 2226

Student Investigator: Natalia Derbentseva
University of Waterloo, Department of Management Sciences
(519) 884-4567 Ext. 2974

You are invited to participate in a research study that investigates the intensity of insight people experience when they solve problems. As a participant in this study, you will be shown a pair of problems at a time alone with their solutions, and you will be asked to rate both problems in each pair on a 10 point scale of difficulty and answer a few questions about the problems. You will be asked to complete 6 such pairs.

Participation in this study is voluntary, and will take approximately half hour of your time. By volunteering for this study, you will learn about research in management sciences in general and the topic of problem solving in particular. In addition, you will receive a summary of the study results when the study is complete (September - October 2006). If you choose to complete this task you will receive **(value of the credit)** toward your final **(course name)** course grade.You may decline to answer any questions presented in this document if you so wish. All information you provide is considered completely confidential; indeed, your name will not be included or in any other way associated, with the data collected in the study. Furthermore, because the interest of this study is in the average responses of the entire group of participants, you will not be identified individually in any way in any written reports of this research. All data collected during this study with identifying information removed will be retained indefinitely, in a locked office to which only researchers associated with this study have access. There are no known or anticipated risks associated with participation in this study.

I would like to assure you that this study has been reviewed and received ethics clearance through the Office of Research Ethics at the University of Waterloo. However, the final decision about participation is yours. If you have any comments or concerns resulting from your participation in this study, please contact Dr. Susan Sykes at this office at (519) 888-4567 Ext. 6005.

If you have any questions about participation in this study, please contact Natalia Derbentseva at (519) 888-4567 ext. 2974 or by email at nderbent@engmail.uwaterloo.ca

Thank you for your assistance in this project.

Please enter your name and student ID below. By entering your name and student ID into this form you give your concent to participate in this study.

First Name

Last Name

StudentID

Page 1 of 17

Instructions

Please, read all the instructions carefully

This is a PDF interactive form, you can complete this form electronically using **Adobe Acrobat Reader (version 6.0.2 or later)** and submit it via e-mail. **You will not be able to save the completed form.**

Before you begin, make sure you are using Adobe Reader version 6.0.2 or later.

You **must have version 6.0.2 or later** of **Adobe Acrobat Reader** to submit this form via e-mail. There is a set of instructions on the last page of this document on how to submit your responses via e-mail.

If you have an older version of Adobe Acrobat Reader, you will need to upgrade your Adobe Reader **before** you can begin filling out this form.

Please **FOLLOW THE INSTRUCTIONS VERY CAREFULLY**. It is very important for this research project that you took this task seriously and did your best in providing your answers.

This task has to be completed in one session, make sure you have at least 30 minutes available before beginning this activity.

In this activity you will be asked to compare 6 pairs of problems. Each pair of problems is set up as a series of steps. It is very important that you complete the steps in the sequence they are described. **Read the instructions for each step at least twice before completing it**. It is very important that you complete each step before proceeding to the next step.

Thank you very much for your participation!

Instructions

On the following pages you will see 6 pairs of problems with their solutions. For each pair, you are asked to judge which of the two problems most people found to be more difficult and which of them most people found to be more insightful when they solved them. We are interested if people can predict the difficulty of a problem without actually solving it but understanding its solution. Your answers will be compared to the responses of individuals who solved these problems in previous studies.

The problems that you will see are called "matchstick arithmetic" problems. A matchstick arithmetic problem represents an unbalanced equation constructed with a set of identical sticks (matchsticks). The goal of the problem is to bring an equation into balance by **moving only one single stick**

E.g.

|| + | = | (2 +1 = 1)

where "II" , "=" and "+" are constructed with 2 sticks
As you can see in these equations numbers are represented as Roman numerals, and mathematical operations are represented as symbols. Below is a chart that clarifies what symbols represent

I = 1	**VII** = 7	"+" is a plus sign
II = 2	**VIII** = 8	"–" is a minus sign
III = 3	**IX** = 9	"=" is an equality sign
IV = 4	**X** = 10	"/ " is a division sign
V = 5	**XI** = 11	"X" is a multiplication sign
VI = 6	**XII** = 12	

A tilted stick, such as a part of "**V**" is not equal to "**I**". Each problem is solved when the equation is brought to balance. This is done by moving only one of the sticks and putting it in a different location within the equation.

E.g.

move
this stick put it here

|| + | = | → | + | = || (1+ 1= 2)

Sticks can not be doubled or completely removed from the equation. Moving a tilted stick "\" or "/" into a straight position "|" also constitutes one move. The problem must be solved by moving only one stick.

Go on to the next page

Instructions (Cont.)

On the following pages you will see 6 pairs of matchstick problems with their solutions. One of your tasks is to judge which of the two problems most people found to be more insightful.

Insight is an "Aha! I see now!" experience, it is an experience of a "light bulb flashing" above your head. As the famous story about Archimedes' discovery goes, while he was taking a bath, he noticed that his body submerged in a bathtub caused the displacement of water from the tub. This led Archimedes right away to discover the principles of density and buoyancy. Extremely delighted by his discovery, Archimedes leapt from his bathtub and ran naked through the streets of Syracuse yelling, "Eureka! Eureka!" ("I've found it! I've found it!"). The Archimedes' story is an example of an extreme level of intensity of an "aha!" experience. You might have had this experience when you suddenly realized how to solve a puzzle or a problem that seemed unsolvable for a while; you had this experience when you "got" a good joke. Most likely at that time you did not act as Archimedes did and your experience of "aha!" was not as intense as his was. All people have this experience with different levels of intensity in various situations.

Below you will see 6 pairs of matchstick problems together with their solutions. The two problems in each pair that you will see might not be equally insightful. Your job is to judge which of the two problems most people found to be more insightful when they solved them. Give your honest best estimate. Before stating your judgment, it is very important that you understand how each problem is solved. **Please take your time to appreciate each of the solutions before making any judgments.**

Go on to the next page

Pair 1

Please read the instructions carefully. Complete **all** the required items on this page before proceeding to the next page.

Step 1: Read the first problem carefully and try solving it yourself for 15 seconds.

Step 2: After 15 sec., carefully examine the solution to the first problem. Make sure you understand which stick was moved and where it was placed.

Step 3: Read the second problem carefully and try solving it yourself for 15 seconds.

Step 4: After 15 sec., carefully examine the solution to the second problem. Make sure you understand which stick was moved and where it was placed.

Problem 1	**Problem 2**
By moving only one stick balance the equation III + III = III	By moving only one stick balance the equation II – I = VI
Solution to Problem 1	**Solution to Problem 2**
III + III = III → III = III = III move this stick put it here	II – I = VI → II – I = I / I move this stick put it here

IV = 4
V = 5
VI = 6
VII = 7
VIII = 8
IX = 9
X = 10
XI = 11
XII = 12

A large group of people participated in a study where they actually had to solve these problems. Your job is to judge which of the two problems most people found to be more difficult and which problem they found to be more insightful. The people who worked on these problems had never solved or seen any other matchstick equations before.

Step 5: On the scales below indicate how most people who solved these problems rated the difficulty of each problem relative to each other

○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Easy				Difficulty of **problem 1**					Difficult
○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Easy				Difficulty of **problem 2**					Difficult

Step 6: On the scales below indicate how most people who solved these problems rated the "insightfulness" of each problem relative to each other

○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Weak insigt				Insightfulness of **problem 1**					Intense insight
○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Weak insigt				Insightfulness of **problem 2**					Intense insight

Pair 1 (Cont.)

Step 7: Please provide any comments that might be helpful in understanding your responses	

Before proceeding to the next page, take a few moments to **refresh your mind**.

When judging the next pair, try to imagine how the next two problems would seem to someone who has never seen this type of problems before.

In other words, do not let your familiarity with previous problems and their solutions to influence your judgment of the next pair.

Try to imagine how it felt to solve these problems for most people who saw matchstick equations for the very first time.

Pair 2

Please read the instructions carefully. Complete **all** the required items on this page before proceeding to the next page.

Step 1: Read the first problem carefully and try solving it yourself for 15 seconds.

Step 2: After 15 sec., carefully examine the solution to the first problem. Make sure you understand which stick was moved and where it was placed.

Step 3: Read the second problem carefully and try solving it yourself for 15 seconds.

Step 4: After 15 sec., carefully examine the solution to the second problem. Make sure you understand which stick was moved and where it was placed.

Problem 1	**Problem 2**
By moving only one stick balance the equation III – II = II	By moving only one stick balance the equation II – I = VI
Solution to Problem 1	**Solution to Problem 2**
III – II = II → II = II = II move this stick put it here	II – I = VI → II – I = I / I move this stick put it here

IV = 4
V = 5
VI = 6
VII = 7
VIII = 8
IX = 9
X = 10
XI = 11
XII = 12

A large group of people participated in a study where they actually had to solve these problems. Your job is to judge which of the two problems most people found to be more difficult and which problem they found to be more insightful. The people who worked on these problems had never solved or seen any other matchstick equations before.

Step 5: On the scales below indicate how most people who solved these problems rated the difficulty of each problem relative to each other

○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Easy				Difficulty of **problem 1**					Difficult
○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Easy				Difficulty of **problem 2**					Difficult

Step 6: On the scales below indicate how most people who solved these problems rated the "insightfulness" of each problem relative to each other

○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Weak insigt				Insightfulness of **problem 1**					Intense insight
○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Weak insigt				Insightfulness of **problem 2**					Intense insight

Pair 2 (Cont.)

Step 8: Please provide any comments that might be helpful in understanding your responses	

Before proceeding to the next page, take a few moments to **refresh your mind**.

When judging the next pair, try to imagine how the next two problems would seem to someone who has never seen this type of problems before.

In other words, do not let your familiarity with previous problems and their solutions to influence your judgment of the next pair.

Try to imagine how it felt to solve these problems for most people who saw matchstick equations for the very first time.

Page 8 of 17

Pair 3

Please read the instructions carefully. Complete **all** the required items on this page before proceeding to the next page.

Step 1: Read the first problem carefully and try solving it yourself for 15 seconds.

Step 2: After 15 sec., carefully examine the solution to the first problem. Make sure you understand which stick was moved and where it was placed.

Step 3: Read the second problem carefully and try solving it yourself for 15 seconds.

Step 4: After 15 sec., carefully examine the solution to the second problem. Make sure you understand which stick was moved and where it was placed.

Problem 1	**Problem 2**
By moving only one stick balance the equation IX = III + I	By moving only one stick balance the equation II – I = VI
Solution to Problem 1	**Solution to Problem 2**
IX = III + I → IV = III + I move this stick put it here	II – I = VI → II – I = I / I move this stick put it here

IV = 4
V = 5
VI = 6
VII = 7
VIII = 8
IX = 9
X = 10
XI = 11
XII = 12

A large group of people participated in a study where they actually had to solve these problems. Your job is to judge which of the two problems most people found to be more difficult and which problem they found to be more insightful. The people who worked on these problems had never solved or seen any other matchstick equations before.

Step 5: On the scales below indicate how most people who solved these problems rated the difficulty of each problem relative to each other

○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Easy				Difficulty of **problem 1**					Difficult
○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Easy				Difficulty of **problem 2**					Difficult

Step 6: On the scales below indicate how most people who solved these problems rated the "insightfulness" of each problem relative to each other

○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Weak insigt				Insightfulness of **problem 1**					Intense insight
○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Weak insigt				Insightfulness of **problem 2**					Intense insight

Pair 3 (Cont.)

Step 8: Please provide any comments that might be helpful in understanding your responses	

Before proceeding to the next page, take a few moments to **refresh your mind**.

When judging the next pair, try to imagine how the next two problems would seem to someone who has never seen this type of problems before.

In other words, do not let your familiarity with previous problems and their solutions to influence your judgment of the next pair.

Try to imagine how it felt to solve these problems for most people who saw matchstick equations for the very first time.

Pair 4

Please read the instructions carefully. Complete **all** the required items on this page before proceeding to the next page.

Step 1: Read the first problem carefully and try solving it yourself for 15 seconds.

Step 2: After 15 sec., carefully examine the solution to the first problem. Make sure you understand which stick was moved and where it was placed.

Step 3: Read the second problem carefully and try solving it yourself for 15 seconds.

Step 4: After 15 sec., carefully examine the solution to the second problem. Make sure you understand which stick was moved and where it was placed.

Problem 1	**Problem 2**
By moving only one stick balance the equation IV – III = III	By moving only one stick balance the equation VII – II = III
Solution to Problem 1	**Solution to Problem 2**
IV – III = III → VI – III = III move this stick put it here	VII – II = III → VI – III = III move this stick put it here

IV = 4
V = 5
VI = 6
VII = 7
VIII = 8
IX = 9
X = 10
XI = 11
XII = 12

A large group of people participated in a study where they actually had to solve these problems. Your job is to judge which of the two problems most people found to be more difficult and which problem they found to be more insightful. The people who worked on these problems had never solved or seen any other matchstick equations before.

Step 5: On the scales below indicate how most people who solved these problems rated the difficulty of each problem relative to each other

○ **1**	○ **2**	○ **3**	○ **4**	○ **5**	○ **6**	○ **7**	○ **8**	○ **9**	○ **10**
Easy				Difficulty of **problem 1**					Difficult
○ **1**	○ **2**	○ **3**	○ **4**	○ **5**	○ **6**	○ **7**	○ **8**	○ **9**	○ **10**
Easy				Difficulty of **problem 2**					Difficult

Step 6: On the scales below indicate how most people who solved these problems rated the "insightfulness" of each problem relative to each other

○ **1**	○ **2**	○ **3**	○ **4**	○ **5**	○ **6**	○ **7**	○ **8**	○ **9**	○ **10**
Weak insigt				Insightfulness of **problem 1**					Intense insight
○ **1**	○ **2**	○ **3**	○ **4**	○ **5**	○ **6**	○ **7**	○ **8**	○ **9**	○ **10**
Weak insigt				Insightfulness of **problem 2**					Intense insight

Pair 4 (Cont.)

Step 8: Please provide any comments that might be helpful in understanding your responses	

Before proceeding to the next page, take a few moments to **refresh your mind**.

When judging the next pair, try to imagine how the next two problems would seem to someone who has never seen this type of problems before.

In other words, do not let your familiarity with previous problems and their solutions to influence your judgment of the next pair.

Try to imagine how it felt to solve these problems for most people who saw matchstick equations for the very first time.

Pair 5

Please read the instructions carefully. Complete **all** the required items on this page before proceeding to the next page.

Step 1: Read the first problem carefully and try solving it yourself for 15 seconds.

Step 2: After 15 sec., carefully examine the solution to the first problem. Make sure you understand which stick was moved and where it was placed.

Step 3: Read the second problem carefully and try solving it yourself for 15 seconds.

Step 4: After 15 sec., carefully examine the solution to the second problem. Make sure you understand which stick was moved and where it was placed.

Problem 1	**Problem 2**
By moving only one stick balance the equation VI + II = III	By moving only one stick balance the equation IV = III – I
Solution to Problem 1	**Solution to Problem 2**
VI + II = III → VI – III = III move this stick put it here	IV = III – I → IV – III = I move this stick put it here

IV = 4
V = 5
VI = 6
VII = 7
VIII = 8
IX = 9
X = 10
XI = 11
XII = 12

A large group of people participated in a study where they actually had to solve these problems. Your job is to judge which of the two problems most people found to be more difficult and which problem they found to be more insightful. The people who worked on these problems had never solved or seen any other matchstick equations before.

Step 5: On the scales below indicate how most people who solved these problems rated the difficulty of each problem relative to each other

○ **1**	○ **2**	○ **3**	○ **4**	○ **5**	○ **6**	○ **7**	○ **8**	○ **9**	○ **10**
Easy				Difficulty of **problem 1**					Difficult
○ **1**	○ **2**	○ **3**	○ **4**	○ **5**	○ **6**	○ **7**	○ **8**	○ **9**	○ **10**
Easy				Difficulty of **problem 2**					Difficult

Step 6: On the scales below indicate how most people who solved these problems rated the "insightfulness" of each problem relative to each other

○ **1**	○ **2**	○ **3**	○ **4**	○ **5**	○ **6**	○ **7**	○ **8**	○ **9**	○ **10**
Weak insigt				Insightfulness of **problem 1**					Intense insight
○ **1**	○ **2**	○ **3**	○ **4**	○ **5**	○ **6**	○ **7**	○ **8**	○ **9**	○ **10**
Weak insigt				Insightfulness of **problem 2**					Intense insight

Pair 5 (Cont.)

Step 8: Please provide any comments that might be helpful in understanding your responses

Before proceeding to the next page, take a few moments to **refresh your mind**.

When judging the next pair, try to imagine how the next two problems would seem to someone who has never seen this type of problems before.

In other words, do not let your familiarity with previous problems and their solutions to influence your judgment of the next pair.

Try to imagine how it felt to solve these problems for most people who saw matchstick equations for the very first time.

Pair 6

Please read the instructions carefully. Complete **all** the required items on this page before proceeding to the next page.

Step 1: Read the first problem carefully and try solving it yourself for 15 seconds.

Step 2: After 15 sec., carefully examine the solution to the first problem. Make sure you understand which stick was moved and where it was placed.

Step 3: Read the second problem carefully and try solving it yourself for 15 seconds.

Step 4: After 15 sec., carefully examine the solution to the second problem. Make sure you understand which stick was moved and where it was placed.

Problem 1	Problem 2
By moving only one stick balance the equation III – II = II	By moving only one stick balance the equation IV = III – I
Solution to Problem 1	**Solution to Problem 2**
III – II = II → II = II = II move this stick put it here	IV = III – I → IV – III = I move this stick put it here

IV = 4
V = 5
VI = 6
VII = 7
VIII = 8
IX = 9
X = 10
XI = 11
XII = 12

A large group of people participated in a study where they actually had to solve these problems. Your job is to judge which of the two problems most people found to be more difficult and which problem they found to be more insightful. The people who worked on these problems had never solved or seen any other matchstick equations before.

Step 5: On the scales below indicate how most people who solved these problems rated the difficulty of each problem relative to each other

○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Easy				Difficulty of **problem 1**					Difficult
○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Easy				Difficulty of **problem 2**					Difficult

Step 6: On the scales below indicate how most people who solved these problems rated the "insightfulness" of each problem relative to each other

○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Weak insigt				Insightfulness of **problem 1**					Intense insight
○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Weak insigt				Insightfulness of **problem 2**					Intense insight

Pair 6 (Cont.)

Step 8: Please provide any comments that might be helpful in understanding your responses	

This is the end of the activity.

Now, go on to the next page to submit your responses by e-mail.

Please follow the instructions on the next page.

Submitting your responses by e-mail.

You must have version 6.0.2 or later of Adobe Acrobat Reader to be able to submit this form via e-mail. To submit your response via e-mail, follow these steps:

1. After you finish completing the form, go over it again to make sure you did not miss any steps.

Note: If you try to e-mail your responses with some missing information, you will receive an error message stating that at least one required field was empty on export. You will need to go back and fill in all missing information.

2. When you finish completing the document you need to submit it by e-mail before exiting. You will not be able to save your answers using Acrobat Reader.

3. At the bottom of this page there is a button "**Submit by e-mail.**" Click that button and follow the prompts to send the email message. The first prompt window will ask you choose how you would like to send this e-mail: using a desktop application, internet e-mail or "other".
If you choose desktop e-mail application, in the next prompt window select "Send data file" and an e-mail message will be automatically created for you with the data file attached.

Note: Adobe will convert your responses into an .XML file automatically.

If you chose one of the other two e-mailing options, follow the steps in the prompt window to save your responses into an .xml file and then you will need to create an e-mail message yourself, attach your saved .xml data file to the e-mail and send it to nderbent@engmail.uwaterloo.ca with an appropriate subject.

Note: depending on the version of your Adobe Reader, you might not be given a choice of how to send your e-mail. The program might automatically open an e-mail message in your desktop e-mail application such as Eudora or Outlook. In this case an e-mail message will be automatically created for you with all the necessary information and your data file attached.

Note: if you clicked "Submit by e-mail" button and no prompts appeared or an e-mail message was not automatically created, it means that your data was not sent successfully. Most likely you have an old version of Adobe Reader

4. Send the e-mail message with the .xml data file attached

5. **Make sure you receive a confirmation e-mail** that your data file was received. If you did not receive a confirmation e-mail within 48 hours of e-mailing your data file, it means that your data file was not received, and you will not get the bonus mark for this activity.

Thank you very much for your participation!

Appendix E: Transcript of verbal instructions given to participants in Experiment 2

Hello! Thank you for volunteering to participate in this study!

My name is Natalia Derbentseva. I'm a PhD student in the department of Management Sciences, and together with Dr. Frank Safayeni I am doing research on human problem solving. This study is a part of data collection that I am doing for my thesis.

With this study we want to get a better understanding of what people experience when they work on problems, both during the solution process and at the very moment of solution itself.

I will ask you to solve two matchstick puzzles and tell me about your experience during the whole process of working and solving these problems. Specifically, I will ask you

- How much stress or tension were you feeling as you were working on the problem,
- How did solving the problem make you feel
- Was there a moment of insight when you solved the problem. By insight I mean a sudden change in the way you see the problem which allows you to realize the solution to the problem right away. This experience of insight is often represented with an "illuminating light bulb" above one's head and an expression of "AHA! I see it now!"
- I will ask you to compare the two problems

I will ask you to solve two matchstick equation puzzles. A matchstick equation puzzle is an equation constructed with Roman numerals and mathematical operations using matchsticks. In the experiment I will use these brown coffee stick to construct these problems for you (show the sticks).

Here is an example of a matchstick equation puzzle (construct on the table the following equation):

II + II = II

It reads "2 + 2 = 2".

As you can see, this equation is not balanced ($2+2 \neq 2$). In matchstick equation puzzles the equation is initially unbalanced and your goal is to bring it to balance by moving only one stick.

Only 1 stick can be moved.
There are some constraints though:

- The stick that was moved must be put back into the equation (you can not take a stick away);
- sticks can not be doubled <show what that means>;
- any form of inequality (≠, <,>, ≤,≥) is not an acceptable solution
- Sticks can be in 3 different orientations: vertical, horizontal, and diagonal.

An acceptable solution to this example could be something like this (perform the move in front of the participant)

II + I = III

To summarize, a matchstick equation is an unbalanced equation represented with Roman numerals and mathematical operations. Your goal is to balance the equation by moving only one stick.

Roman Numerals

The symbols that can be used in these puzzles are Roman numerals from 1 to 12 and 5 mathematical operations: addition, subtraction, multiplication, division, and equality. There is a list of Roman numerals from 1 to 12 with their corresponding Arabic equivalents on the table. You can use this list any time during the experiment.

Now I will show you all these symbols constructed with sticks. (construct and show all 12 Roman numerals and 5 operations>

Would you like to review the Roman numerals again? Do you have any questions?

To summarize, matchstick equation puzzles are unbalanced equations represented with Roman Numerals and mathematical operations. Your goal is to balance the equation by moving only one stick. You can not use any form of inequality. You can not remove or double the sticks.

During the experiment I will ask you to solve two of these problems. The problems will be constructed with coffee sticks and laid out inside this folder in front of you. You will be working on one problem at a time and after you solve each problem I will ask you several questions about your experience solving this problem.

Different pairs of problems are used in this experiment. In some pairs one problem is easier than the other, and in other pairs both problems are relatively similar in their difficulty. Which pair you'll get will be determined randomly and I will ask you to judge the relative difficulty of the two problems after you solve them.

Let's begin?

Questions to the participants for providing judgments

For the first problem

- Could you please draw a graph that represents how your feeling of tension or stress was changing (if it was) while you were working on the problem and after you solved it, until now?

- <<if the graph showed a decrease in tension after the solution>> You indicated that your tension decreased after the solution, how long did take for your tension to decrease from solution to the lowest level you indicted after the solution

- When you found the solution to this problem did you have a Eureka moment? It is a moment when light bulb flashes above one's head?

- <<if the answer was affirmative>> Could you please go back in your mind to that specific moment when you realized the solution to this problem and "live through it" one more time, so it stays fresh in your memory. You do not need to describe it to me, just for you to keep it in mind

- <<before proceeding to the next problem>> could you please go back in your mind to your solution moment on the first problem and get a good feel for it one more time.

For the second problem

- Could you please draw a graph of your feeling of tension or stress on the second problem? Please show me with your drawing how this feeling on the second problem compares t the first (i.e. on which problem you experienced more tension at different intervals). Please use the distance between the curves to show the relative difference in your feeling.

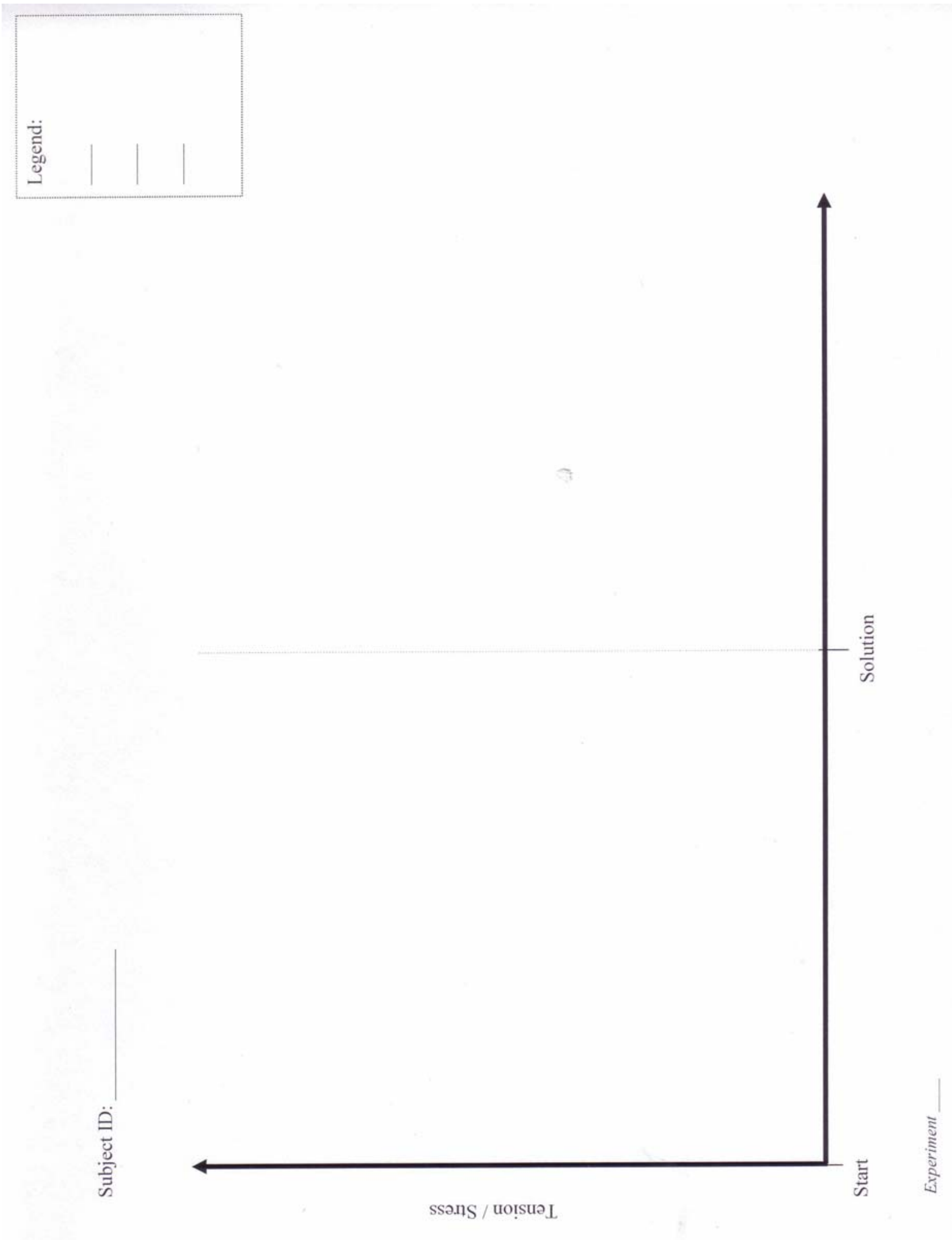
- <<if the graph showed a decrease in tension after the solution>> You indicated that your tension decreased after the solution, how long did take for your tension to decrease from solution to the lowest level you indicted after the solution

- Did you have a Eureka moment on the second problem?

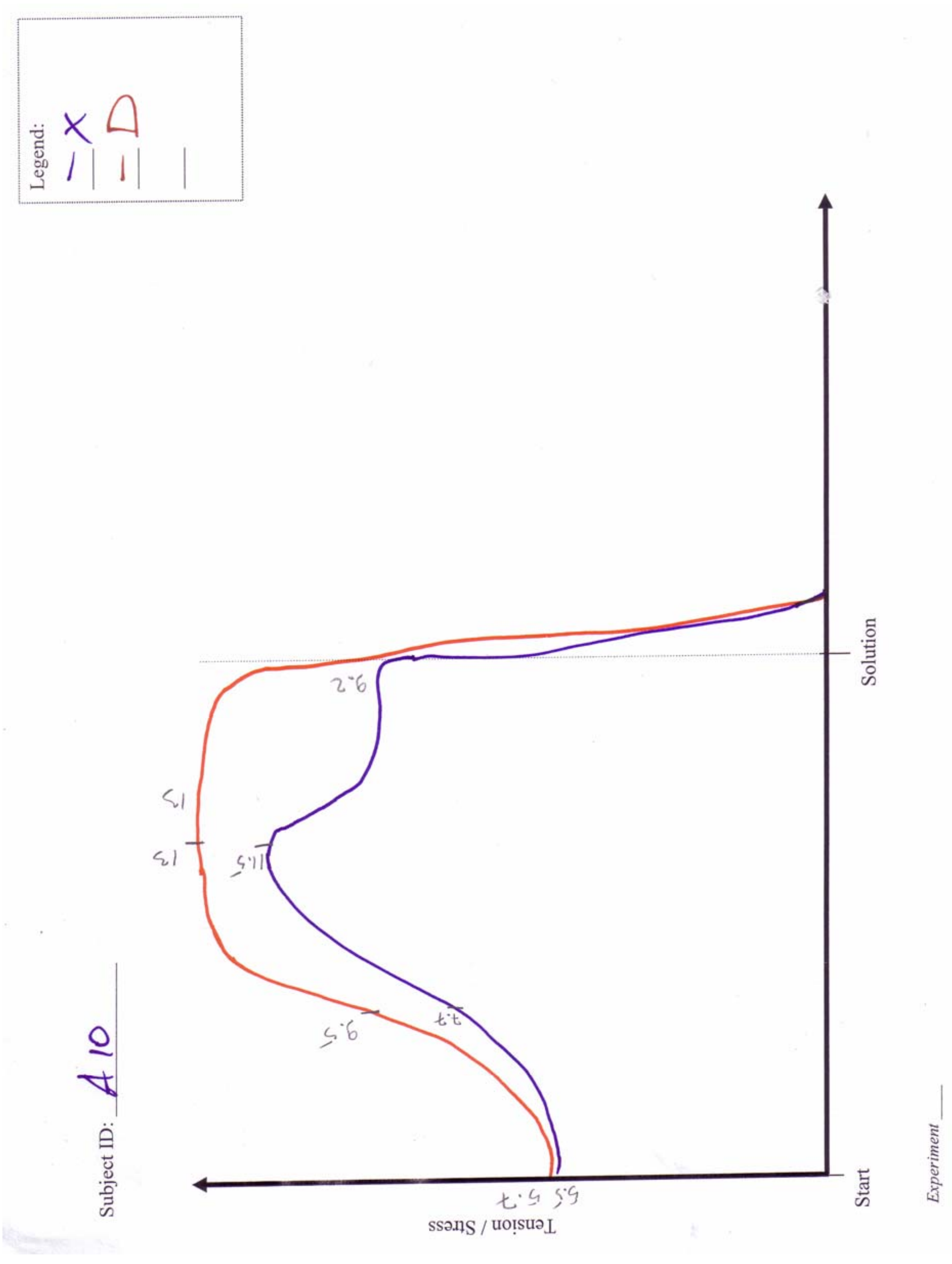
- <<if the answer was affirmative, both problems were constructed on the table and their solutions were shown to the participant one at a time, and the participant was asked to "get a good feel" for the moment when they realized that solution. Both problems were presented like this>> On which of these two problem was your Eureka more intense? Could you please indicate the difference in intensity of the two Eureka experiences on this scale? Please put two points on this scale, representing your solution moments, and space these points apart on this scale such that the distance between the points, relative to the distance from the start, represents the difference in the intensity of the Eureka moments that you experienced.

- Could you please indicate how these two problems relate to each other in terms of their difficulty? Please put the two problems on this scale positioning them such that the distance between them on this scale gives a rough idea of the difference in difficulty based on the way you experienced these two problems.

Appendix F: Blank tension graph used in Experiments 2, 3, and 4 (reduced size)



Appendix G: Example of a completed tension graph (reduced size)



Appendix H: Transcript of video instructions used in Experiments 3 and 4

General Introduction Part

Hello!
My name is Natalia Derbentseva. I'm a PhD student in the department of Management Sciences, and together with Dr. Frank Safayeni I am doing research on human problem solving. The study that you volunteered to participate in is a part of data collection that I am doing for my thesis.

Thank you for participating in this study!

In my research I've tried to get a better understanding of what we experience when we work on problems, what happens to us when we solve a problem, how we experience a moment at which we realize the solution to a problem.

To get your help on that, I will ask you to solve two puzzles and tell me about your experience during the whole process of working and solving these problems. Specifically, I will ask you

- How much stress or tension were you feeling as you were working on the problem,
- How did solving the problem make you feel
- Was there a moment of insight when you solved the problem. By insight I mean a sudden change in the way you see the problem which allows you to realize the solution to the problem right away. This experience of insight is often represented with an "illuminating light bulb" above one's head and an expression of "AHA! I see it now!"
- I will ask you to compare the two problems

To get you more familiar with the specific questions that I will ask during the study, we will go through 2 practice problems before we start the actual experiment.

The structure of the study will be the following:

- First, we will go through 2 practice problems to give you an idea of what kind of questions I will ask you after you solve the experimental problems
- After that, I will introduce you to the problems that you will be working on in the study
- Then, you will solve two of those problems and answer the questions about your experience on those problems. And that will conclude our session.

Now, let's start with 2 practice problems. Remember, that this is just a training, and its purpose is to give you an idea of what kind of information I will ask you to provide when you solve the real experimental problems

During this training you'll be working on problems called "anagrams." In the actual experiment you'll be working on different type of problems. As you might know, an anagram is a string of letters that can be rearranged using each letter exactly once to produce other word(s).

As a problem to solve in this training, I will give you a word and your job will be to rearrange its letters to make a different word. You can use each letter from the initial word only once. To make this training similar to the actual experiment, I will ask you to solve two of these anagram problems. You will be working on them one at a time, and after you solve each one of them I will ask you several questions about your experience on these problems the same questions as I will ask you during the actual experiment.

Remember, that this is just a training, and its purpose is to give you an idea of what kind of questions I will ask you after you solve the real experimental problems

Thank you.

Roman numerals and Matchstick equation puzzles training part

Now I will introduce you to the problems that you'll be working on in the actual experiment. These problems are called matchstick equation puzzles. A matchstick equation puzzle is an equation constructed with Roman numerals and mathematical operations using matchsticks. In the experiment I will use brown coffee stick to construct these problems for you.
Here is an example of a matchstick equation puzzle:

<The equation below was shown on the screen>

II + II = II

It reads "2 + 2 = 2".

As you can see, this equation is not balanced ($2+2 \neq 2$). In matchstick equation puzzles the equation is initially unbalanced and your goal is to bring it to balance by moving only one stick.

Only 1 stick can be moved.
There are some constraints though:

- The stick that was moved must be put back into the equation (you can not take a stick away);
- sticks can not be doubled <show what that means>;
- any form of inequality ($\neq$, <,>, $\leq$,$\geq$) is not an acceptable solution

- Sticks can be in 3 different orientations: vertical, horizontal, and diagonal.

An acceptable solution to this particular example could be something like this

<the solution is demonstrated>

II + I = III

To summarize, a matchstick equation is an unbalanced equation represented with Roman numerals and mathematical operations. Your goal is to balance the equation by moving only one stick.

Roman Numerals

The symbols that can be used in these puzzles are Roman numerals from 1 to 12 and 5 mathematical operations: addition, subtraction, multiplication, division, and equality. There is a list of Roman numerals from 1 to 12 with their corresponding Arabic equivalents on the table. You can use this list any time during the experiment.

Now I will show you all these symbols constructed with sticks.

<construct and show on the screen all 12 Roman numerals and 5 operations>

To summarize, matchstick equation puzzles are unbalanced equations represented with Roman Numerals and mathematical operations. Your goal is to balance the equation by moving only one stick. You can not use any form of inequality. You can not remove or double the sticks.

During the experiment you will be asked to solve two of these problems. The problems will be constructed with coffee sticks and laid out on the table in front of you. You will be working on one problem at a time and after you solve each problem I will ask you several questions about your experience solving this problem.

Different pairs of problems are used in this experiment. In some pairs one problem is easier than the other, and in other pairs both problems are relatively similar in their difficulty. Which pair you'll get will be determined randomly and I will ask you to judge the relative difficulty of the two problems after you solve them.

Good luck with your problems!

Thank you very much!

Questions to the participants for providing judgments

For the first problem

- Could you please draw a graph that represents how your feeling of tension or stress was changing (if it was) while you were working on the problem and after you solved it, until now?

- <<if the graph showed a decrease in tension after the solution>> You indicated that your tension decreased after the solution, how long did take for your tension to decrease from solution to the lowest level you indicted after the solution

- When you found the solution to this problem did you have a Eureka moment? It is a moment when light bulb flashes above one's head?

- <<if the answer was affirmative>> Could you please go back in your mind to that specific moment when you realized the solution to this problem and "live through it" one more time, so it stays fresh in your memory. You do not need to describe it to me, just for you to keep it in mind

- <<before proceeding to the next problem>> could you please go back in your mind to your solution moment on the first problem and get a good feel for it one more time.

For the second problem

- Could you please draw a graph of your feeling of tension or stress on the second problem? Please show me with your drawing how this feeling on the second problem compares t the first (i.e. on which problem you experienced more tension at different intervals). Please use the distance between the curves to show the relative difference in your feeling.

- <<if the graph showed a decrease in tension after the solution>> You indicated that your tension decreased after the solution, how long did take for your tension to decrease from solution to the lowest level you indicted after the solution

- Did you have a Eureka moment on the second problem?

- <<if the answer was affirmative, both problems were constructed on the table and their solutions were shown to the participant one at a time, and the participant was asked to "get a good feel" for the moment when they realized that solution. Both problems were presented like this>> On which of these two problem was your Eureka more intense? Could you please indicate the difference in intensity of the two Eureka experiences on this scale? Please put two points on this scale, representing your solution moments, and space these points apart on this scale such that the distance between the points, relative to the distance from the start, represents the difference in the intensity of the Eureka moments that you experienced.

- Could you please indicate how these two problems relate to each other in terms of their difficulty? Please put the two problems on this scale positioning them such that the distance between them on this scale gives a rough idea of the difference in difficulty based on the way you experienced these two problems.